Rationally designed high-temperature polymer dielectrics for capacitive energy storage: An experimental and computational alliance

Pritish S Aklujkar [a,1], Rishi Gurnani [b,f,1], Pragati Rout [c], Ashish R Khomane [a], Irene Mutegi [c], Mohak Desai [a], Amy Pollock [a], John M Toribio [c], Jing Hao [d], Yang Cao [d,e,*], Rampi Ramprasad [b,f,*], Gregory Sotzing [a,c,*]

[a] Institute of Materials Science, University of Connecticut, Storrs, CT 06279, USA
[b] School of Materials Science and Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[c] Department of Chemistry, University of Connecticut, Storrs, CT 06269, USA
[d] Electrical Insulation Research Center, Institute of Materials Science, University of Connecticut, Storrs, CT 06269, USA
[e] Department of Electrical and Computer Engineering, University of Connecticut, Storrs, CT 06269, USA
[f] Matmerize Inc., Atlanta, GA 30308, USA

ARTICLE INFO

ABSTRACT

Polymer-based electrostatic capacitors find critical use in high-temperature applications such as electrified aircraft, automobiles, space exploration, geothermal/nuclear power plants, wind pitch control, and pulsed power systems. However, existing commercial all-organic polymer dielectrics suffer from significant degradation and failure at elevated temperatures due to their limited thermal stability. Consequently, these capacitors require additional cooling systems, that require increased system load and costs. Traditionally, an inability to directly predict or model key properties - such as thermal stability, breakdown strength, and energy density has been an impediment to the design of such polymers. To enhance the experimentation and instinctive-driven approach to polymer discovery there has been recent progress in developing a modern co-design approach. This review highlights the advancements in a synergistic rational co-design approach for all-organic polymer dielectrics that combines artificial intelligence (AI), experimental synthesis, and electrical characterization. A particular focus is given to the identification of polymer structural parameters that improve the capacitive energy storage performance. Important structural elements, also known as proxies, are recognized with the rational co-design approach. The central constituents of AI and their influence on accelerating the discovery of new proxies, and polymers are presented in detail. Recent success and critical next steps in the field showcase the potential of the co-design approach.

1. Introduction

Grid distribution networks are progressively facing more and more challenges due to the increasing need for electric vehicle (EV) charging stations, the growing adoption of renewable energy sources, the requirement for effective energy storage solutions, and the significant impacts of climate change [1]. The changing dynamics of electricity sources, loads, and consumption patterns highlight the imperative for improved scalability, distribution, and flexibility in their efficient administration. The advancement of energy storage technology is a crucial factor for the stability, consistency, and high-field performance of energy systems. Energy storage components – electrostatic and electrochemical capacitors, batteries, and fuel cells are gaining increased attention. In comparison to energy storage devices which utilize electrochemical reactions to store energy, dielectric capacitors, which store electrical energy through an electrostatic field, are distinguished by their exceptional power densities attributable to their swift charge-discharge capabilities within brief time frames [2–4]. Polymer-based dielectrics overcome the disadvantages of ceramic dielectrics by offering a comparatively higher breakdown strength, flexibility, ease of processing, variable structural design, and a smoother failure mode [5,6]. Electrostatic (dielectric) capacitors in particular offer high power density, high operating voltage, and lesser loss as compared to other classes of energy storage components [7]. For applications involving extreme thermal and electrical conditions, all-organic polymer-based electrostatic capacitors are vital in achieving high energy

* Corresponding authors.
E-mail addresses: yang.cao@uconn.edu (Y. Cao),
rampi.ramprasad@mse.gatech.edu (R. Ramprasad), g.sotzing@uconn.edu (G. Sotzing).
[1] These authors contributed equally to this work.

Nomenclature

AI	Artificial intelligence
BNNS	Boron nitride nano sheets
BOPP	Biaxially oriented polypropylene
DFT	Density functional theory
ECFP	Extended connectivity fingerprint
E_{at}	Atomization energy
E_{bd}	Dielectric breakdown field
e_{coh}	Cohesive energy density
Egap	Bandgap
EV	Electric vehicle
FVEs	Free volume elements
GA	Genetic algorithm
GNN	Graph neural network
GPR	Gaussian process regression
K	Dielectric constant
KRR	Kernel ridge regression
LLM	Large language models
MD	Machine data
MGI	Material genome initiative
ML	Machine learning
MLP	Multilayer perceptron
NLP	Neural language processing
PDTC-HDA	Poly para-phenylene diisothiocyanate hexane diamine
PDTC-HK511	Poly para-phenylene diisothiocyanate jeffamine
PDTC-MDA	Poly para-phenylene diisothiocyanate diphenylmethanediamine
PDTC-ODA	Poly para-phenylene diisothiocyanate oxydianiline
PDTC-PhDA	Poly para-phenylene diisothiocyanate phenylmethanediamine
PEEK	Polyether ether ketone
PEI	Polyetherimide
PEN	Polyethylene naphthalate
PET	Polyethylene terephthalate
PG	Polymer genome
PI	Polyimide
PMDA	Pyro metallic gianhydride
PNB-2Me5Cl	2-methyl-5-choro polynorbornene
PNB-3Cl4Me	3-chloro-4-methyl polynorbornene
PNB-2,5DM	Dimethyl polynorbornene
POFNBs	Polyoxafluoronorbornenes
PolyG2G	Polymer graph-to-graph
PONB-2Me5Cl	2-methyl-5-chloro polyoxanorbornene
PVK	Polyvinylkarbazole
ROAM	Restricted orientation anisotropy method
RPD	Probability distribution
SD-VAE	Syntex directed variational autoencoder
SMILES	Simplified molecular input line entry system
Tg	Glass transition temperatures
Tm	Melting temperature
U_d	Energy density
ε_r	Relative permittivity
ε_o	Vacuum permittivity
ε_{elec}	Electronic component of the dielectric constant
ε_{ion}	Ionic component of the dielectric constant
ϕ_e^{Al}	Electron injection barrier

density, and optimum efficiency for harsh condition electrifications.

For linear dielectrics, the induced polarization is proportional to the total external and internal electric fields. For such dielectric, the energy density (U_d) can be derived [3] and represented as shown:

$$U_d = 1/2\varepsilon_o\varepsilon_r E^2 \tag{1}$$

As seen from Eq. (1), the energy density varies with the square of the applied electric field (E, often interpreted as the breakdown strength), and the dielectric constant (ε_r, relative permittivity) varies linearly with the energy density. The charging and discharging energy density along with the above-mentioned variables play an important role in tuning the charge-discharge efficiency of the polymer dielectrics [8].

For high-temperature applications, the electronic systems are usually exposed to temperatures above 150 °C, thus dielectric polymers should be stable and efficient for energy storage at such high temperatures. Polyolefins commonly have a high bandgap and low conduction loss which is preferred for high electric field applications [9]. However, these polymers cannot withstand high temperatures due to their low glass transition temperatures (Tg). Generally, the application temperatures for such polymer dielectrics are usually <100 °C. For example, biaxially oriented polypropylene (BOPP), which is a widely used polymer dielectric has a high breakdown strength of ≈700 MV m^{-1} and small loss, but can be operated efficiently to a maximum operating temperature of only 85 °C [10,11]. Such polymer dielectrics require an external cooling system which boosts their cost and required maintenance limiting their scope of application. Polymers with conjugated aromatic backbones having a high Tg have been reported for high-temperature energy storage applications, but these polymers display a higher conduction loss due to their lower bandgap which further leads to a drop in electric field endurance and an amplified energy loss. Low thermal stability not only limits the maximum operating temperature of all-organic polymer dielectrics but also limits the maximum energy density of these dielectrics is the low dielectric constant (K) of polymers [4]. Extensive research in improving the dielectric constant involved adding fillers in the polymer matrix. Although these fillers enhance the overall permittivity of the polymer matrix by providing more polarizable sites, they cause a significant deterioration in the breakdown strength as the amount of the filler exceeds the percolation threshold [3]. The deterioration of the local electric field is mainly due to the big difference in dielectric constant between the filler and the polymer. Generally, the dielectric constant of nanocomposites will have a significant increase when the ratio of fillers is close to the percolation threshold. This can lead to a current pathway and thus a reduction of breakdown strength. For example, fillers like BNNS have been used to improve the breakdown strength and reduce conduction losses owing to their high band gap (5.5–5.8 eV) and excellent thermal stability. However, it was seen that the breakdown strength increased up to 10 % to 12 % BNNS loading followed by a reduction of the breakdown strength as the % loading increased [12,13].

Several previous studies have demonstrated the significance of bandgap, dielectric constant, free volume, and high breakdown strength on polymer dielectric performance [14]. Deciphering structural elements in the polymer backbone and their effect on a particular dielectric property is fundamental for the discovery of new polymer dielectrics. This article presents insights into the polymer structural elements and discusses their dependence on the dielectric properties for high-temperature and high-density energy storage. Considering the mutual constraints between the properties and the scope for extensive structural modification of polymer chains, a co-design approach that combines experimental investigation and computational (AI-based) techniques is highly valuable. Elements of artificial intelligence, machine learning, and
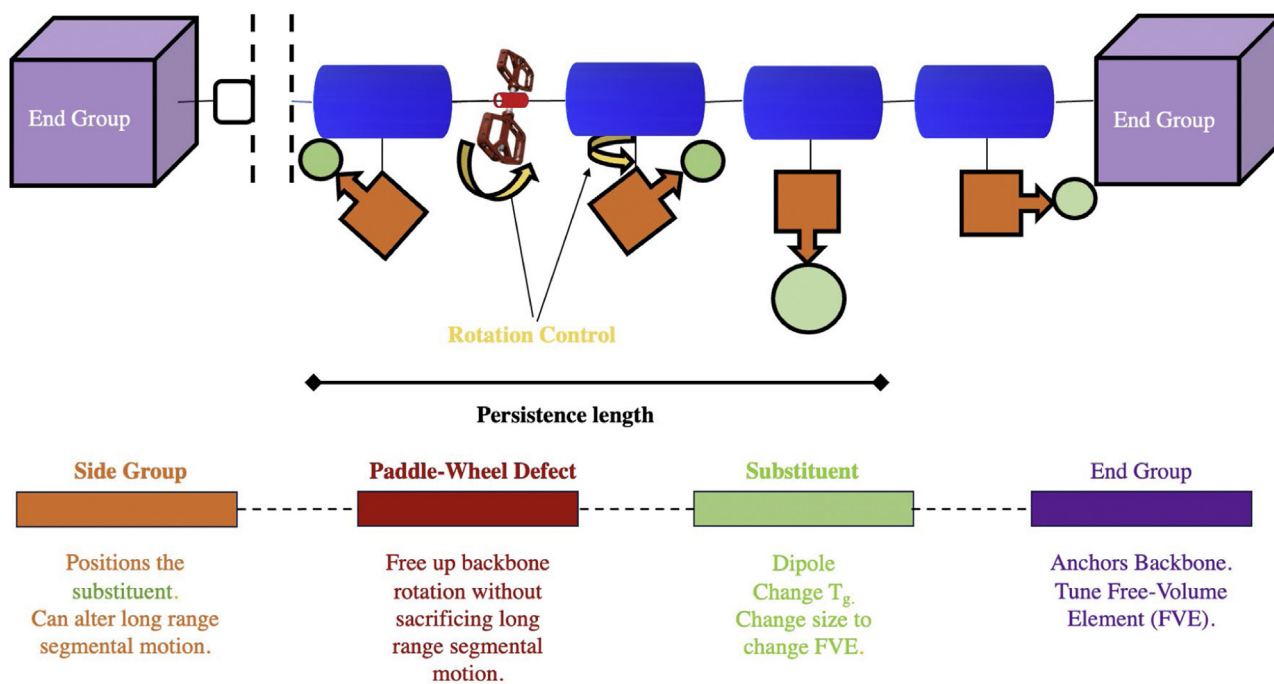
Fig. 1. Polymer chain modifications using several proxies for optimal capacitive energy storage.

high-throughput computational screening of dielectric properties are explored further which can accelerate the discovery of new polymer dielectrics.

2. Rational Co-design strategy for all-organic polymer dielectrics

Given the complexity involved in polymer structure design and dielectric performance, a sophisticated rational approach is necessary. In the pursuit of developing novel all-organic polymer dielectrics, a primary objective is to devise a strategy for discovery that aims to achieve polymers with specific properties optimized for capacitive energy storage. These properties are intricately linked to various structural elements within the polymer chain, referred to as proxies, through structure-property relationships. Numerous proxies can be discerned from existing literature, and polymer chain structures can be designed accordingly. An example of polymer chain design is shown in Fig. 1 where different proxies are highlighted. The impact of different proxies on the capacitive energy storage performance is further discussed in detail (Section 3).

The conventional approach for polymer discovery requires synthesizing the polymer experimentally and assessing its properties through experimental measurements. Based on these results, adjustments may be made to some proxies, leading to the synthesis of new polymers and subsequent experimental measurements. This conventional method is time-consuming and resource-intensive, particularly for all-organic polymer dielectrics, where multiple structure-property relationships are involved, necessitating elaborate synthesis, processing, and characterization procedures. To mitigate the drawbacks of the conventional method, the discovery strategy can be enhanced through the integration of a co-design methodology, combining computational techniques with experimental measurements. A co-design methodology fosters a symbiotic relationship between experimental and computational approaches, wherein each complements the other. This facilitates the acceleration of polymer discovery while reducing costs and time expenditures. Fig. 2 highlights the major steps involved in all-

organic polymer dielectrics discovery. This approach proves advantageous in uncovering novel proxies, thereby broadening the scope for the discovery of new polymer dielectrics [15,16].

3. Structure-property insights from proxies for all-organic polymer dielectrics

3.1. Thermal stability

The operating temperature requirements for dielectric capacitors in various electronic systems are shown in Fig. 3. Capacitive energy storage is crucial for space exploration, oil and gas operations, electric vehicles, commercial aircraft, geothermal and nuclear power plants, as well as electrification systems that require stability at high temperatures. The maximum operating temperatures of several commodity polymers generally do not meet the required range of operation and thus cooling systems are coupled with the dielectric capacitor assembly to protect and operate the capacitors. To meet the required demand at high temperatures it is indispensable that the polymers should have inherent high-temperature thermal stability [9].

The two main transition temperatures in polymers are the glass transition temperature (T_g) and the melting temperature (T_m). For amorphous polymers, the thermal stability is determined by T_g, whereas for semi-crystalline polymers it is determined by T_m [17]. At temperature below T_g, the segmental motion of polymer chains is terminated and there is a restricted bond rotation and chain immobility causing the polymer to withhold the structure. As the temperature increases above T_g the chains gain mobility with a concurrent increase in free volume. In such a molten state there is an increase in the leakage current due to a surge in charges and dipolar relaxation which eventually results in a decline of breakdown strength and a rise in polarization loss [18]. Introducing more amorphous segments such as rigid aromatics and bulky groups helps in increasing T_g. Polymers with high crystallinity have a low T_g which inhibits their performance at high temperatures. Highly crystalline polymers are not preferred for high-temperature dielectric applications. One of our pioneering studies on polyimides
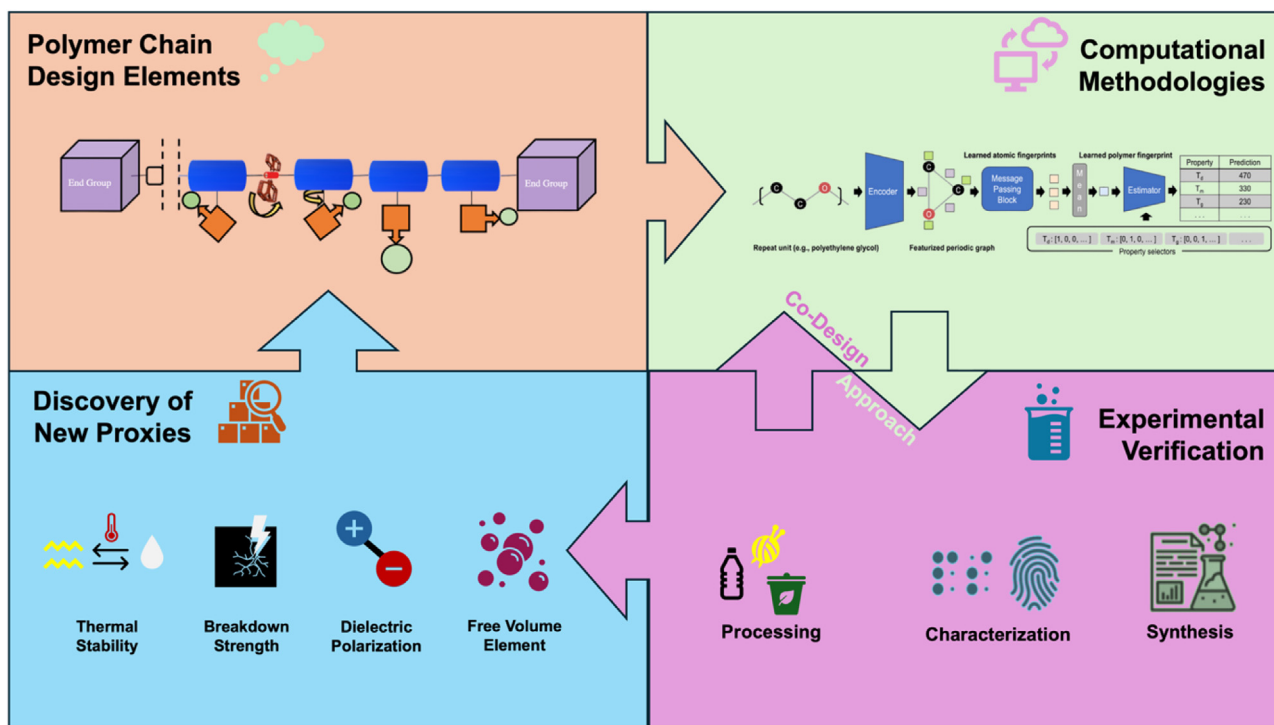
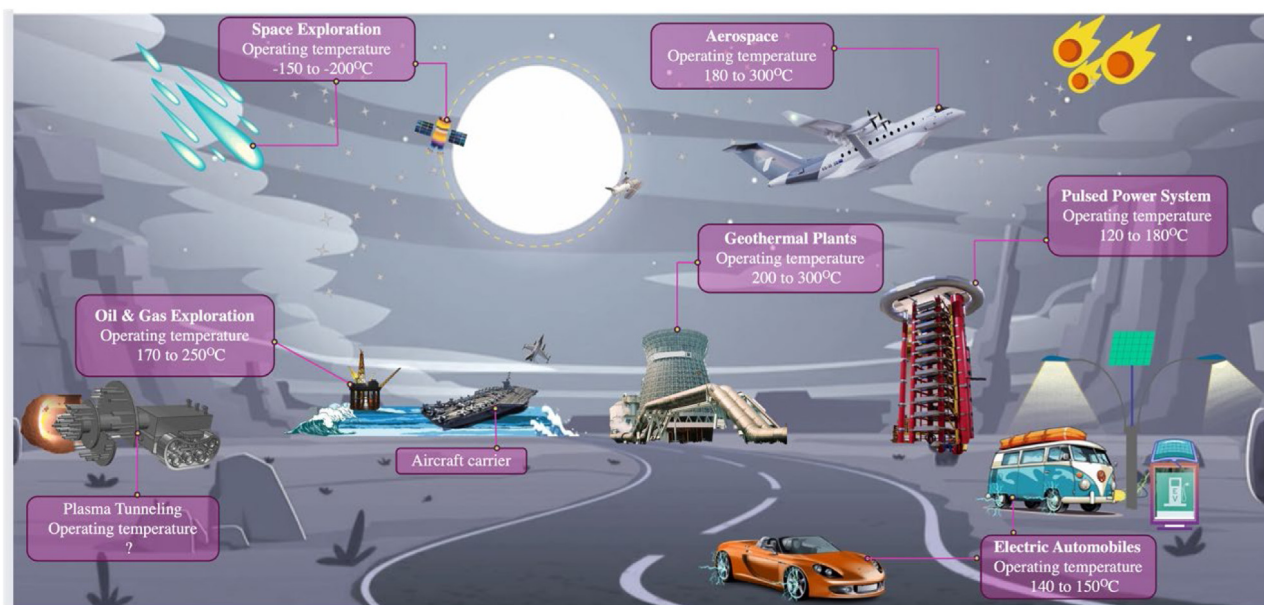**Fig. 2.** The overview of the key elements involved in the rational co-design approach.

Fig. 3. Application areas of high-temperature electrostatic capacitors and respective operation temperatures.

where pyromellitic dianhydride (PMDA) was reacted with several short-chain diamines revealed that when heating above Tg the dielectric loss increases rapidly [19]. Although the synthesized polyimides had a significantly large dielectric constant the Tg was considered a main proxy for failure at high temperatures. Thus, a higher Tg of the polymer is a decisive factor for the performance of dielectric polymer at high temperatures.

Other techniques used to enhance thermal stability are crosslinking, introduction of high thermal conductivity fillers, blends, and fabricating sandwiched polymer structures [3]. Although these modifications increase the thermal stability of the polymers, they continue to face several drawbacks at high tem-

peratures. For example, at high temperatures, crosslinking can result in restricted dipolar motions which results in a lowering of the dielectric constant, and the addition of fillers can display a rise in electrical conductivity which increases the leakage current [14]. At high voltages, the all-organic polymers with high thermal conductivity can dissipate heat better and prevent the thermal breakdown process. Modifications that enhance the thermal conductivity of the polymer also support the thermal stability of polymers with naturally high Tg, which is essential for maintaining thermal stability [20,21].

Several reported high-temperature polymer dielectrics mainly comprise amorphous aromatic chain units. These polymers contain
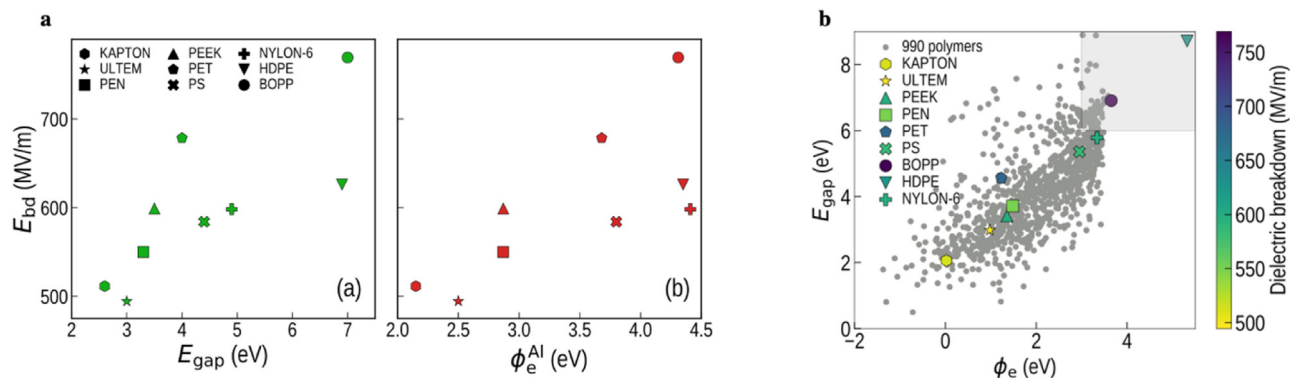
Fig. 4. a E_{bd} for 10 Al-polymer interfaces shown as a function of Egap and $\phi_e{}^{Al}$. b Screened polymers with computed Egap and ϕ_e. (The proxies of 10 polymers named in the legend were overlayed to come up with screening criteria to discover high E_{bd} polymers. The shaded regions are the predicted polymers) [34]. Copyright 2020, Reproduced with permission from the American Chemical Society.

conjugated π bonds which are central to the thermal stability of the polymer as they help in the delocalization of electrons and decrease the energy of the molecular units [8]. In the case of dielectric applications, this phenomenon of delocalization causes a drawback resulting in increased leakage current at higher electric fields. Thus, many polymers with conjugated backbone have a higher Tg but they tend to have lower bandgap due to increased conductivity at higher temperatures. Having a higher Tg is one of the many conditions required for high-temperature polymer dielectric. Thus, a rational co-design approach is a must when designing polymers with high Tg without conceding the bandgap [14].

3.2. Breakdown strength

Dielectrics are divided primarily into organic and inorganic materials. Among the several requirements for all-organic polymer dielectrics, the breakdown field strength is extremely important for a material to be used in high-field dielectric applications. Organic materials are preferred over inorganic materials due to their higher flexibility and graceful breakdown strengths [22–25]. Dielectric breakdown field (E_{bd}) is defined as the value of the strongest external electric field a material can withstand before becoming electrically conductive, leading to the breakdown of the material [26,27].

Polymer dielectric breakdown can be categorized into intrinsic breakdown and extrinsic breakdown. The intrinsic breakdown process is related to the electronic structure of pure dielectric materials. The extrinsic breakdown process is related to the accumulation of structural defects and charge carriers over a long time resulting from aging and field-induced material degradation [28]. The interplay of thermal, electronic, and mechanical properties together impacts the dielectric breakdown of a polymer. This has resulted in a lack of complete understanding of the factors or mechanisms leading to the breakdown of materials [29–32]. Apart from the Artbauer theory of dielectric breakdown, it has also been stated that as high-density current is applied breakdown occurs when the polymer rapidly loses its ability to resist the passage of current through it [33].

Given the complexity of the dielectric breakdown mechanism, it is a difficult task to evaluate structural design for polymers with high breakdown strength. Our efforts in the field have revealed the importance of various proxies on which the breakdown strength of polymer dielectrics depends. An innovative approach was further adopted in screening potential polymers that can have high breakdown strength E_{bd}. As computing the E_{bd} is demanding we investigated other interdependent properties and aimed to find correlations that can enhance the E_{bd}. It was found that the E_{bd} of the polymer when interfacing with a metal electrode correlates with

the bandgap (Egap) and the electron injection barrier ($\phi_e{}^{Al}$) at the interface, Fig. 4a shows the direct proportionality of the two proxies with the breakdown strength [34]. Using such correlations and computational methods to validate the proxies we were able to screen 990 previously made polymers and discover 53 potential polymers which can exhibit high E_{bd} (Fig. 4b).

An overview of our computational and experimental studies exposes the dependence of three proxies namely, bandgap, charge injection barrier, and cohesive energy density. These proxies are of primary importance in enhancing the breakdown strength of polymers. Dielectric breakdown mechanisms are intricate and primarily determined by factors such as material composition, electrical conductivity, thermal conductivity, film thickness, operating temperature, and electrical stress conditions [35]. Not limited to these proxies other parameters can affect the breakdown strength of polymers. For instance, the gradual decline in capacitive storage performance of polymer dielectrics at high temperatures is linked to the rise in leakage current caused by thermal and electric fields [36–39]. The rise in conduction as the electrical resistivity of polymers decreases with applied field and temperature is well known [39]. Structural modification such as introducing a local rotation in the polymer backbone is a useful proxy to mitigate conduction loss and increase the chain flexibility. The improved chain flexibility improves the bandgap and prohibits localized defect states. Polymers with inherently large bandgaps are more likely to exhibit higher breakdown strength and lower conduction loss.

To explore the impact of the rotational barrier on the conduction loss, we modified the polyetherimide (PEI) with structural defects (paddle-wheel effect) (Fig. 5a and b) by introducing a small amount of p-phenylenediamine which has a lower rotational energy barrier around the N (imide) – C (aromatic) bond for the polymer as compared to the conventional m-phenylenediamine which was confirmed using comparative results [40,41] as shown in Fig. 5c. The introduction of 5 % p-phenylenediamine (5p-PEI) increased the bandgap of modified PEI to 3.32 eV compared to the bandgap of PEI which was 3.24 eV. PEI is one of the outstanding commercial polymers which has a high Tg. However, at elevated temperatures, it displays a rise in electric conduction hampering the energy storage performance at high temperatures. A minor modification in polymer structure design significantly affected the high-temperature energy storage performance without majorly affecting the Tg of the polymer. Modified PEI due to its enhanced flexibility displayed a better charge-discharge efficiency and repressed conduction loss at high temperatures as shown in Fig. 5d. The maximum electric fields for 5p-PEI were statistically higher than PEI as shown in Fig. 5e indicating a higher breakdown strength.
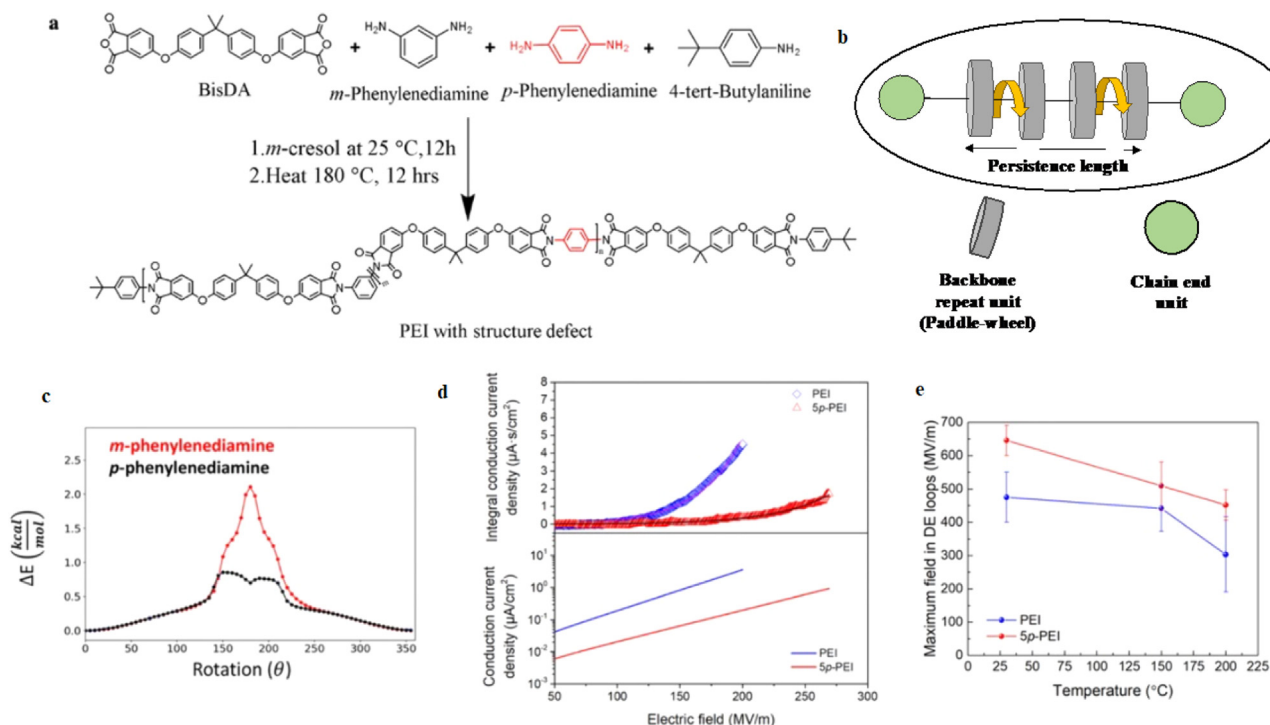
Fig. 5. a Synthesis of PEI with the structural defect. b Schematic representation of paddle-wheel effect with repeating units having rotational freedom. c Rotational energy barrier observed for the PEI chains. d High electric field conduction of PEI and modified PEI (5p- indicating 5 % defect) at 150 °C. e The maximum electric field in DE loops for PEI and 5p-PEI as a function of the temperature. (The data points are average values and error bars denote the standard deviation) [40,41]. Copyright 2020, and 2022, Reproduced with permission from the Institute of Electrical and Electronics Engineers, and American Chemical Society.

3.3. Dielectric polarizations

Polarization can be expressed as the sum of all the dipole moments per unit volume. Dipole moments are essential for the dielectric performance of the polymers. Polarization (P) can be related to dielectric permittivity (ε_r) and the applied electric field (E) according to Eq. (2) [42]. Dielectric permittivity (ε_r) is a frequency-reliant complex number that accounts for the developed polarizations in an insulator in response to an external electric field. The real part of ε_r is related to the energy storage capability and the imaginary part determines the dielectric loss [43]. In an applied electric field, the more the polarizations developed in the polymer more the dielectric constant. Also, from Eq. (1) the dielectric permittivity (dielectric constant) is directly proportional to the energy density (U_d) of the capacitor. Hence, polarizations, or a high dielectric constant are ideal for the high efficiency and high energy density performance of the capacitor.

$$P = E\varepsilon_o(\varepsilon_r - 1) \qquad (2)$$

Depending on the source of dipoles, the polarizations can be distinguished into three types – electronic, atomic, and dipolar polarization. All these polarizations depend on the frequency of the electric field and influence the total permittivity. As shown in Fig. 6, the electronic and atomic polarizations occur at a higher frequency, as these involve displacement of electrons and alterations in atomic positions respectively which requires higher energy. At lower frequencies (operating frequencies for dielectric capacitors) the dipolar polarizations, which are a result of dipolar rearrangements with the electric field, have a major effect on the capacitor performance. The relaxation of these polarizations results in dielectric loss which leads to energy dissipation over the charging and discharging cycles. The polarization losses due to electronic and atomic polarization are insignificant in the range of operation.
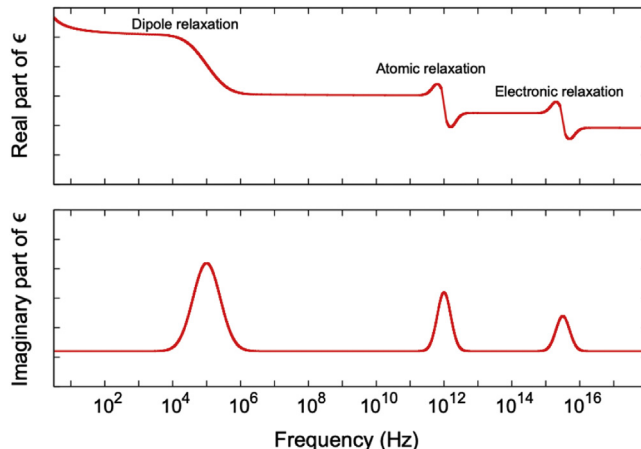

Fig. 6. Schematic of the real and imaginary parts of the orientationally averaged dielectric permittivity [43]. Copyright 2016, Reproduced with permission from Elsevier Science Ltd.

Whereas the loss due to dipolar relaxations plays a major role in capacitive energy storage.

High temperature and high electric fields have an inverse effect on the dipole movements. The high temperature increases the thermal motions which hinder the alignment of the dipoles in response to the external field and eventually lowers the polarization. Therefore, the polymers should have a constant dielectric constant over a range of temperatures, which is generally not the case for low Tg polymers as the dielectric constant gets altered as the temperature bumps above the Tg. Several tactics to improve dielectric polarizations such as including more dipoles in the polymer back-

bone, insertion of high ε_r fillings into the polymer matrix, and enhancing the flexibility of the dipoles have been studied over the past. As many of these alterations increase the dielectric loss care must be taken in adjusting the degree of these modifications. For example, when the amount of the conducting fillers is near the percolation threshold the dielectric losses due to conduction increase drastically. Similarly, the introduction of excess dipoles can lead to long-range dipolar interaction which can further result in tragic dielectric loss at higher temperatures.

We initially investigated to explore the dielectric performance of polymers by enhancing the dielectric constant. The structure-property relationship was developed for polyurea and polyurethane thin films with dielectric constant against the dielectric loss [44]. It was found that the dielectric constant increases and dielectric loss decreases as the number of carbons between polarizable function groups decreases. This was supported by the data obtained by the high throughput density functional theory (DFT) calculations. In a subsequent study, the influence of different proxies including dielectric constant and bandgap on capacitive energy storage were examined [45]. DFT calculations were used to perform an initial screening and down selection of polymer structures. Through this rational co-design, polyimides were determined to be an important class material for high dielectric constant and low loss for energy storage applications. One of the rationally co-designed polyimides was able to achieve the highest dielectric constant 7.8 with a potential energy density of ~15J/cm^3. This study established the importance of the increased dipole volume from the ether linkage as well as the longer conjugation length when a carbonyl spacer was inserted between the benzene ring on the dielectric constant.

With the advancement of computational and experimental analysis to uncover the relationship between the dielectric constant and bandgap it was found that the electronic component of the dielectric constant (ε_{elec}) had an inverse relationship with the Eg. Whereas the ionic component of the dielectric constant (ε_{ion}) and the total non-electronic component has no relation with Eg (Fig. 7a). This further motivated us to develop metal-containing polymers via which ε_{ion} can be enhanced without sacrificing the Eg. The validity of the computational calculations was confirmed with experimental validations, Fig. 7b shows measured and computed similar results. Further exploration of organometallic polymer dielectrics which were predicted to have a higher Eg and high ε (Fig. 7c) led us to develop several metal-containing polymers that had high dielectric constant between 5 and 8 and which also maintained a high bandgap Eg > 6eV [46–57].

For commercial dielectric polymers, it is observed that the dielectric constant and the bandgap follow an inverse relationship (Fig. 7f). As both these parameters are essential for efficient dielectric performance, a cautious approach is needed when increasing the dielectric constant without affecting the bandgap [14]. The complex interdependence of these parameters can be confronted using a rational co-design approach where the design of the polymer structure restricts the dipolar relaxation loss and promotes unrestricted rotation of dipoles [14]. A class of rationally co-designed polyoxafluoronorbornenes (POFNBs) breaks the inverse relation between the dielectric constant and bandgap. The polymer design elements in POFNBs promote strong dipolar relaxation without compromising the bandgap (Fig. 7d, e, and f) [58]. The correlation between the Tg, bandgap, and the dielectric constant is difficult to comprehend when designing polymers. More importantly, the discovery of new proxies like the free volume element, and the paddle-wheel effect, increases the design complications. In such situations, a co-design approach where an AI-based informatics approach can assist the experimental observations would provide key insights into improving and augmenting the polymer design strategy.

3.4. Free volume elements

Free volume in polymers is primarily responsible for several macroscopic properties such as thermal stability, viscosity, gas permeability, and transport properties. It also affects the dielectric constant [61] and the Tg of the polymer. At temperatures below Tg, the polymer chain movement is restricted, making the polymer chains randomly pack with minimal segmental motion. This inefficient packing leaves multiple vacancies of unoccupied space which have different shapes and sizes. These pockets with a radius on the angstrom scale are referred to as the free volume elements (FVEs) [62–65].

FVEs can be used to determine and explain the motion of liquids and solids. They are only a portion of the total percent of the unoccupied volume of a polymer and partially determine the structural heterogeneity of the polymeric materials. Free-volume elements have potentially been proven to influence the electrical properties of all-organic polymer dielectrics [66]. As temperature increases, the polymer experiences different physical conformations due to mobility because of the increase in the free volume. Free volume in a polymer is observed to a greater extent on the polymer chain end than in the units within the chain [67] as the end chains have a longer intermolecular distance with the neighboring chain ends. Free volume also has been shown to have a direct effect on conduction, with various studies and theoretical models predicting increased electron mobility through free space [68–71].

Based on this ideology, recent research was done to compare cyclic polystyrene films compared to their linear chain melt [72]. The authors demonstrated that cyclic polystyrene showed an enhanced dielectric strength and capacitive energy density compared to its linear chain melt which is attributed to the enhanced packing of cyclic polymers due to lack of free chain-ends. Consequently, the void spaces created due to free volume elements in polymers give room for the acceleration of electrons at high fields leading to high dielectric breakdown as demonstrated by Artbauer in the free volume theory. The work confirmed that polymer topology can substantially influence the capacitive properties of all-organic polymer films. A recent collaborative work on free-volume elements in polyetherimide polymers was conducted [67] to show the relationship between the free-volume elements and the breakdown strength. This was achieved by keeping molecular weight, and glass transition temperature constant and only changing the end-capping groups (Fig. 8a). Substituting the end groups is a better strategy than any other structural modification for altering the distribution of free-volume elements. The measurement of free volume elements radius probability distribution curves was done using an ultrafast infrared laser technique known as the Restricted Orientation Anisotropy Method (ROAM). The observations indicate that large FVEs lower the breakdown strength of the polymer leading to polymer breakdown at lower electric fields when FVE increases. The results emphasize the importance of chain packing and restricted free rotations of molecules in space which is often neglected when designing dielectric polymers. Excellent results are achieved when the probability for large FVEs is low as displayed in Fig. 8b. The Gaussian line shape is an indication of low FVEs and correspondingly high E$_{bd}$. This study motivates researchers to consider the effect of voids on dielectric performance as space itself can play a major role in conduction loss and dielectric loss.

4. Accelerated polymer discovery with rational Co-design

While commercial high-temperature all-organic polymer dielectrics offer enhanced thermal stability, their aromatic structures reduce the bandgap, leading to high conduction loss and low breakdown strength. Previous studies have demonstrated the sig-
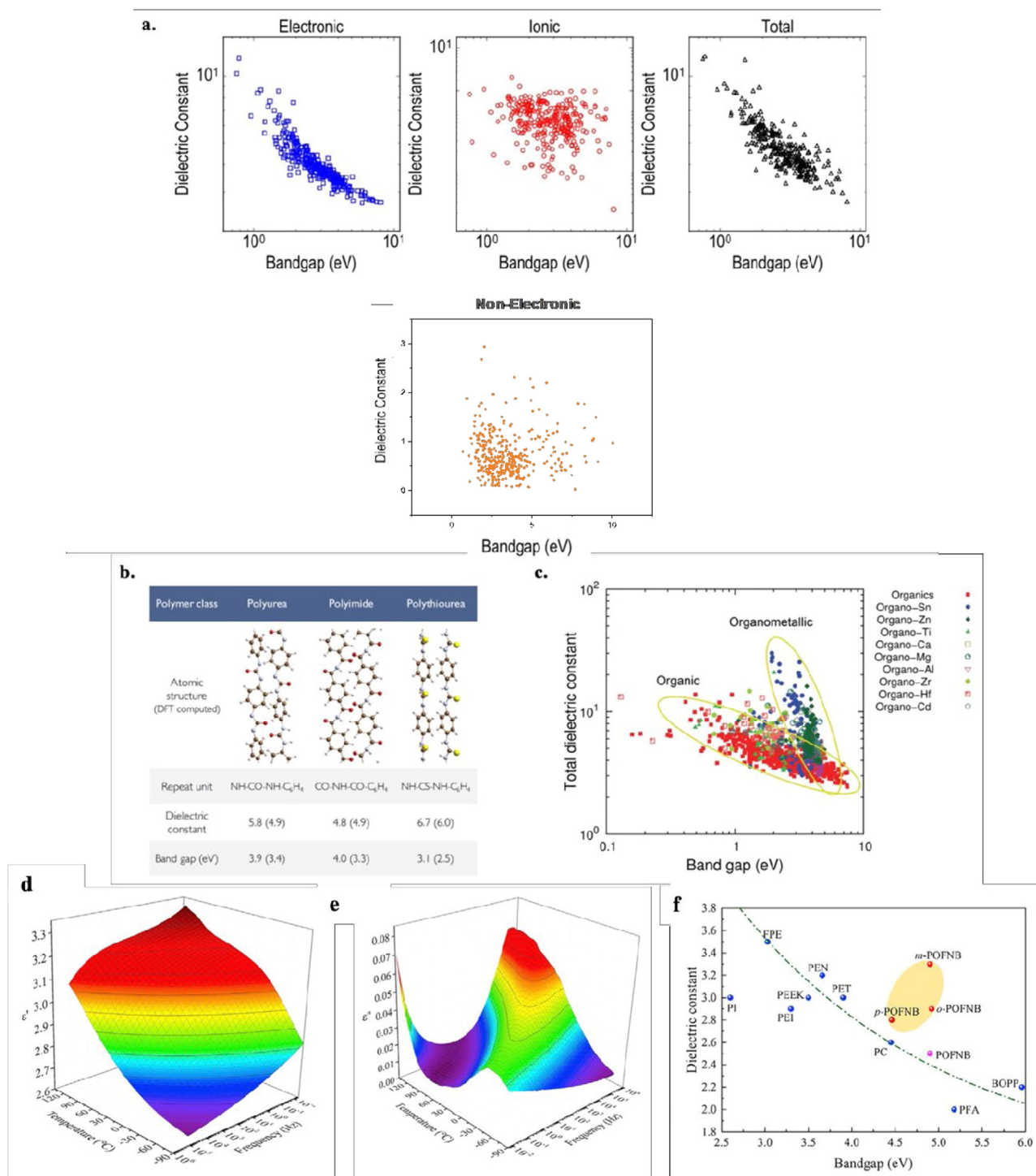
Fig. 7. a Computed bandgap and dielectric constant of the 4-block polymers. b DFT computed structures for the repeat units and the measured and computed (in brackets) values of the first generation of rationally co-designed polymers. c Computational dielectric constant of over a thousand organic and organometallic materials vs their computational bandgap. (the two ovals indicate different spaces which consist of either organic or organometallic materials). d, and e The real part (d) and imaginary part (e) of the relative permittivity of m-POFNB. f Correlation between the measured dielectric constant and band gap for POFNBs and other commercial polymer dielectrics [42,58–60]. Copyright 2018, 2016, 2022, and 2021, Reprinted with permission from Elsevier Science Ltd, John Wiley & Sons Inc, The Royal Society of Chemistry, and Proceedings of the National Academy of Sciences of the United States of America.

nificance of bandgap in polymer dielectric performance. Furthermore, introducing rigidity and thermally conductive fillers does not significantly enhance capacitive energy storage. Polymer structure modifications should yield adequate dipolar relaxation to enhance the dielectric constant without compromising thermal stability or the polymer's bandgap [14].

The interdependence of Tg, ε_r, and Eg on the breakdown strength (E_b) and the energy density (U_d) of polymer dielectric can be explained using the trends shown in Fig. 9a, and b [4]. Both E_b and U_d demonstrate comparable trends at both room temperatures and high temperatures, but a contrasting trend is noticed in the case of ε_r. The fall in the energy density and breakdown strength
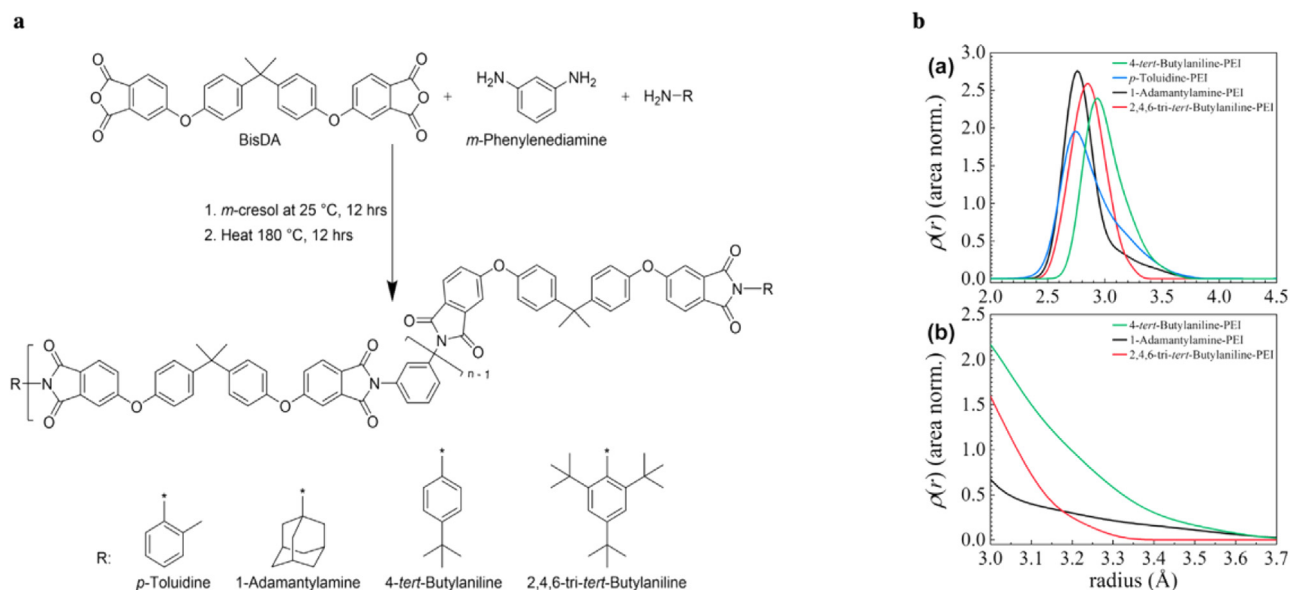
Fig. 8. a Synthesis scheme for polyetherimide (PEI) with different end caps. b FVE radius probability distribution (RPD) in four end-capped PEI films - (a) Probability curves are area normalized to emphasize the variations in line shapes. RPD for 2,4,6 -tri–tert-butylaniline-PEI is similar to a Gaussian curve, and others are non-Gaussian with elongated tails and bigger radii. (b) Zoomed-in representation of (a) where the amplitude of tails is correlated with the breakdown field, the higher the amplitude lower the breakdown strength [67]. Copyright 2023, Reproduced with permission from Elsevier Science Ltd.
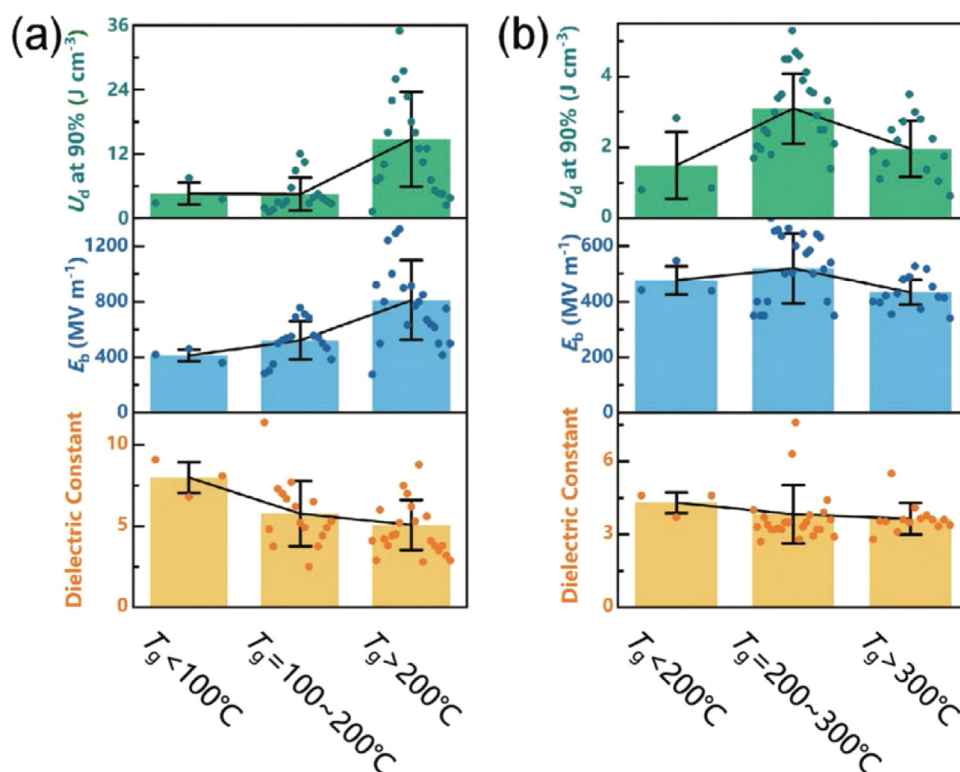
Fig. 9. Difference in energy storage performance of linear polymer dielectrics at a) 25 °C and at b) 150 °C [4]. Copyright 2022, Reproduced with permission from John Wiley & Sons Inc.

for high Tg polymers at high temperatures underscores the challenge of avoiding conduction losses which are due to the presence of the inherent aromatic structure of these polymers. Thus, comparing Fig. 9a and b one gets an idea of how a high operating temperature can alter the dielectric performance of a polymer. This motivates the search for modern rational co-design strategies that can develop polymer structures without compromising all the essential parameters that are responsible for optimum performance.

An example of an effective rational co-design strategy is illustrated by the development of a unique class of polyolefins known as polyoxafluoronorbornenes (POFNBs). These polymers have a structure comprising flexible and rigid bicyclic chemical components that maintain a high Tg and high bandgap (Fig. 10) [11]. The dipolar relaxations are preserved, resulting in a favorable dielectric constant that can be adjusted through the incorporation of different functional groups (Fig. 10c). These polyolefins exhibit
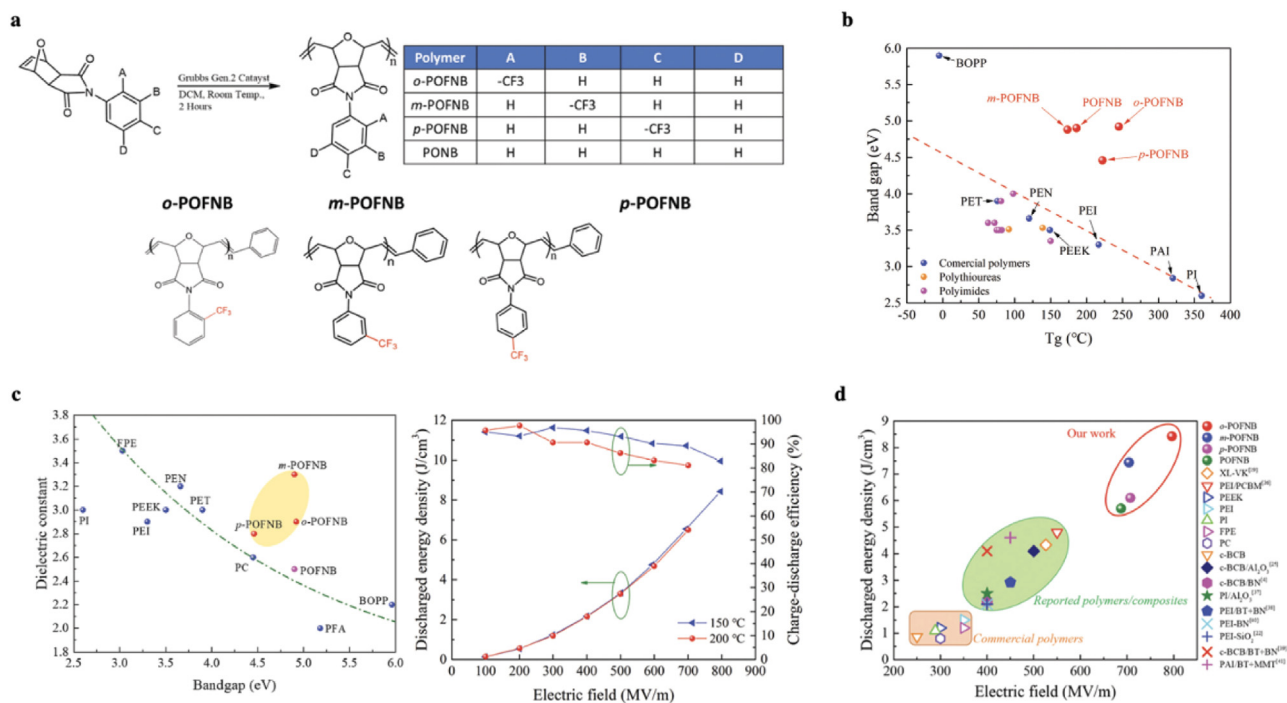
Fig. 10. Structural design and performance of POFNBs. a Polymer structure study. b The bandgap vs. Tg for polymers synthesized in this study and for established polymers with a high electric field and/or high-temperature stability. c Dielectric constants as a function of the bandgap for POFNBs and established all-organic dielectric polymers and energy storage performance of o-POFNB at 150 °C and 200 °C. d Discharge energy density and electric field endurance for POFNBs vs other polymer dielectrics at 150 °C [60]. Copyright 2022, Reproduced with permission from The Royal Society of Chemistry.

Table 1
Comparison of experimental glass transition temperature (Tg), dielectric constant (K), and bandgap (Eg) of commercial polymer dielectrics and rationally designed polymer dielectrics for capacitive energy storage.

Polymers		Tg (°C)	K (r.t 1 kHz)	Eg (eV)	Ref
Commercial	PET	75	3.3	3.9	[46,73]
Polymer	BOPP	−5	2.2	5.9	[46,73]
Dielectrics	PEI	217	2.9	3.3	[42,74]
	PI	360	3	2.6	[42,74]
	PEEK	149	3	3.5	[42,74]
	PEN	120	3.2	3.6	[46,73]
Rationally	PONB-2Me5Cl	232	3.29	4.39	[75]
Designed	PNB-2Me5Cl	243	3.06	4.32	[75]
Polymer	PNB-3Cl4Me	220	2.9	4.27	[75]
Dielectrics	PNB-2,5DM	232	3.14	4.34	[75]
	o-POFNB	245	2.90	4.92	[60]
	m-POFNB	178	3.25	4.84	[60]
	p-POFNB	220	2.80	4.45	[60]
	PDTC-ODA	N/O	4.52	3.22	[54]
	PDTC-MDA	N/O	4.08	3.16	[54]
	PDTC-PhDA	N/O	4.89	3.07	[54]
	PDTC-HDA	139	3.67	3.53	[54]
	PDTC-HK511	92	6.09	3.51	[54]

a high discharge density, outperforming several commercial all-organic polymer dielectrics. Additionally, they have also demonstrated low conduction loss at elevated temperatures. Among the various POFNBs, ortho-POFNB stands out with the highest Tg of 244 °C (Fig. 10b) and an energy density of 6.5 J/cm3 at 200 °C (Fig. 10c). The inherent flexibility of the polymer chain proves crucial in dissipating loss and maintaining stability at high temperatures, providing greater rotational freedom for the polymer chain. The effectiveness of the rational co-design strategy is demonstrated in Fig. 10, showcasing experimental results of various POFNBs at elevated temperatures [60]. A comparative property analysis of commercial polymer dielectrics and selected rationally co-designed polymers is shown in Table 1. The selected polymers are from the

class of rational-designed functional polynorbornenes that show enhanced thermal stability and polythioureas with high dielectric constant. It can be observed that the newly designed polymers break the mutual constraints between the properties discussed in previous sections.

Addressing the interdependence of several proxies and understanding their influence on dielectric properties requires a systematic exploration of polymer space [3,14]. For polymer dielectrics, DFT has proven to be reliable and accurate in predicting several electronic properties and further assisting in the development of several new polymer structures. Fig. 11 shows a few rationally co-designed polymers using DFT which outperform BOPP. However, due to the computational overhead of DFT, only a limited
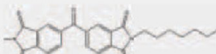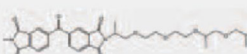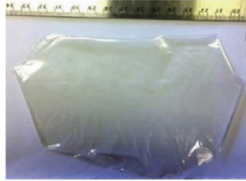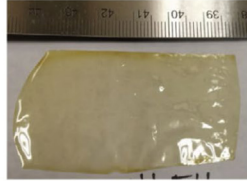
Polymer name	BOPP	PDTC-HDA (Polythiourea)	BTDA-HDA (Polyimide)	BTDA-HK511 (Polyimide)
Repeat unit				
Synthesized polymer	(Metallized)			
Dielectric constant	2.2	3.7	3.6	7.8
Breakdown field (MV/m)	700	685	812	676
Energy density (J/cm3)	~5	~9	~10	~16

Fig. 11. Comparison of rationally co-designed polymers with commercial polymer dielectric BOPP using DFT [59]. Copyright 2018, Reproduced with permission from Elsevier Science Ltd.
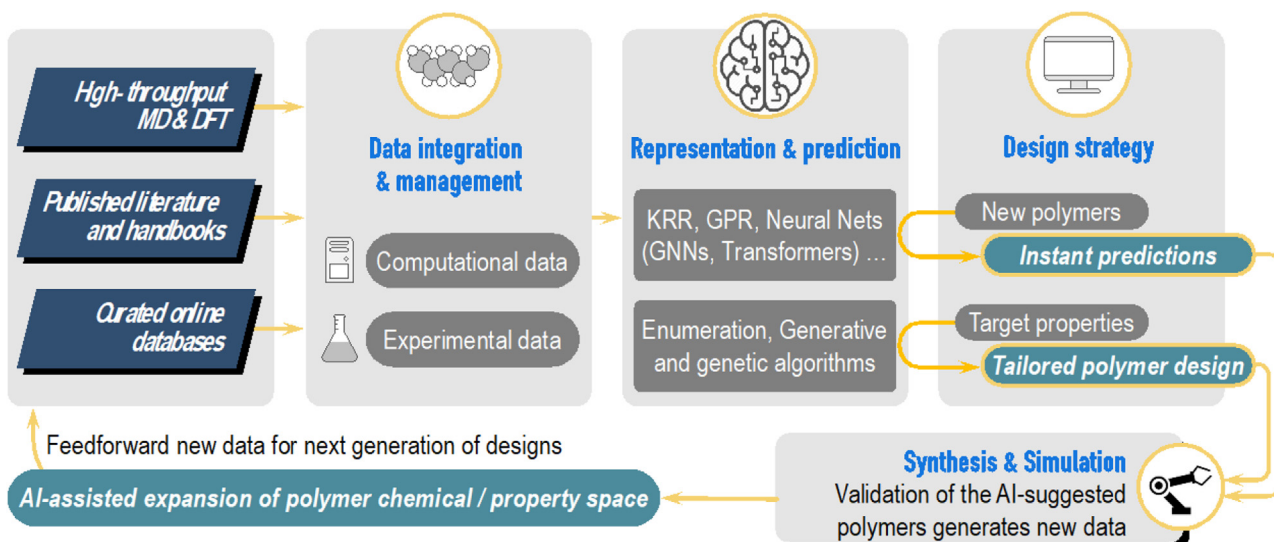
Fig. 12. The ecosystem of AI-based design of polymer dielectrics. The starting point is high-throughput simulations, publications, handbooks, and web sources from which computational and experimental data are integrated and managed. From there, AI algorithms are used in two modes. In the "forward" mode, the algorithms are trained to make instant property predictions for an enumerated set of polymers. In the "inverse" mode, the algorithms are trained to generate polymers on-demand for an enumerated set of target properties. The polymers suggested by either mode undergo validation by physics-based computational methods, chemical synthesis, and physical experimentation; a process that generates new data. That data is then fed back into the loop.

number of polymer chemistries and topologies can be explored. Also, Tg, and therefore thermal stability, cannot be computed using DFT. This is why, for high-temperature dielectric design, AI and ML techniques have been pursued, as they can estimate the Tg by training on large experimental data sets.

5. Elements of artificial intelligence (AI) for high-temperature polymer dielectric co-design

Within the vast expanse of chemical possibilities for polymers, a wide variety of high-performance dielectrics likely await discovery. Well-trained and calibrated artificial intelligence (AI), capable of handling large numbers that challenge human imagination, can help converge on extraordinary or "outlier" materials rapidly [76]. While the AI methods may come in different flavors, most share the common elements listed in Fig. 12: [77] data, representation (i.e., fingerprinting), prediction, and design. These elements will be discussed one by one in the following sections. Table 2 classifies

the primary computational methods that have influenced the design of polymer dielectrics, with a particular emphasis on methods related to artificial intelligence.

5.1. Data

The starting point of any machine learning model is data. In the context of polymer dielectrics, an individual data point takes the form of the tuple [Polymer, Property]—where "Property" corresponds to one of the proxy properties identified in Section 3. This type of data can be acquired from various sources, including handbooks, online repositories, and published journal articles (see Table 2 for examples).

Experimental data extracted from several scientific resources forms a reliable basis of information as it has already been peer-reviewed by the scientific community [78–80]. Several of these data sources have been digitalized enabling users to access the data over different online platforms [81,82]. For example, one of

Table 2

Notable data sources and methods for AI-based design of polymer dielectrics.

Name	Category	Description	Web Source
Prediction of polymer properties, CRC Press [88]	Experimental data set	Handbook of polymer properties including dielectric constant and glass-transition temperature	
Polymer Data Handbook, Oxford University Press [89]	Experimental data set	Handbook of polymer properties including dielectric constant	
PoLyInfo [80]	Experimental data set	~500k point database of polymer properties including glass transition temperature, thermal decomposition temperature, melting temperature, and thermal conductivity	https://polymers.nims.go.jp
Khazana [84]	Computational data set	Database of DFT-computed properties including dielectric constant and band gap	https://khazana.gatech.edu
RandonPy [86]	Computational data set	Database of MD-computed properties including thermal conductivity and dielectric constant	https://github.com/RadonPy/RadonPy
PSP [90]	Enabling technology	A tool to generate the 3D structure of polymers from SMILES strings at various scales (oligomers, loops, crystals, amorphous structures, etc.), thereby expediting high-throughput MD data generation.	https://github.com/Ramprasad-Group/PSP
ECFP [91]	Representation	An algorithm for solving molecular isomorphism has since been adapted into a polymer representation. Not invariant to addition, subtraction, and translation of polymer repeat units.	https://github.com/rdkit/rdkit
PG [92]	Representation	A hierarchical algorithm for representing polymers. Invariant to addition, subtraction, and translation.	https://polymergenome.org/
KRR	General ML algorithm	An algorithm that has been used to develop structure-property models. Limited capacity to model non-linear relationships.	
GPR	General ML algorithm	An algorithm that has been used to develop structure-property models. Can model non-linear relationships. Scales poorly as training data grows.	
polyGNN [93]	Representation, polymer-specific ML algorithm	A graph neural network-based algorithm for representing polymers as graphs. Invariant to addition, subtraction, and translation. polyGNN has also been used to develop structure-property models that scale favorably as training data grows.	https://github.com/Ramprasad-Group/polygnn
polyBERT [94]	Representation, polymer-specific ML algorithm	A Transformer-based algorithm for representing polymers as a language. polyBERT has also been used to develop structure-property models that scale favorably as training data grows.	https://github.com/Ramprasad-Group/polyBERT
Multitask learning [95,96]	Enabling technology	A concept that improves structure-property model accuracy, especially in data-scarce situations.	
Genetic Algorithm [97,98]	General ML algorithm	An algorithm that has been used to generate polymer dielectric chemical structures.	https://github.com/Ramprasad-Group/polyga
SD-VAE [99]	General ML algorithm	A generative algorithm that has been used to generate polymer dielectric chemical structures.	
polyG2G [100]	Polymer-specific ML algorithm	A generative algorithm that casts chemical structure generation as a translation problem over graphs. Has been used to generate polymer dielectrics.	

the pioneering databases "PoLyInfo" has a compilation of several hundreds of polymer structures and properties that can be used for a variety of applications [83]. Given the possible variations in different polymer structures and the vastness of available scientific data, a manual approach to data extraction from literature is usually a strenuous task.

Computations, based on first-principle theory and classical molecular dynamics techniques, offer potential prospects for the automated generation of polymer datasets (albeit at lower fidelity compared to experimental data) [84]. Relatively large computational data sets include Khazana [84,85]which contains polymer band gap and dielectric constant data from DFT calculations, and RadonPy, [86] which contains data on dielectrics constant and a

handful of other properties from MD simulations of various polymers.

MD simulations require an initial 3D atomic structure, and the structure that is chosen can have a significant impact on the trajectory of the simulations and therefore on the calculated materials properties. Traditionally, the creation of this structure has been a manual step and an impediment to scaling up simulation protocols to many different polymers. To address this problem, PSP was developed [85]. The primary input to PSP is the SMILES string of the polymer repeat unit. Optional inputs include polymer chain length, density, size of simulation box for amorphous models, etc. The tool then uses this information, and geometry optimization routines, to generate 3D structures of and loop oligomers, infinite
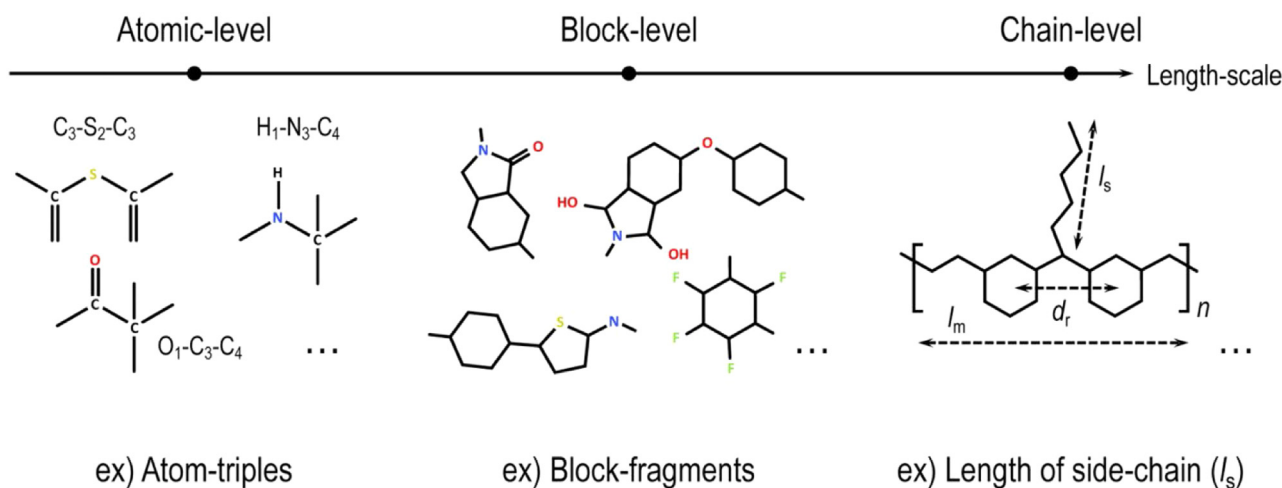
Fig. 13. Hierarchical fingerprints used to represent polymers in the Polymer Genome pipeline [92]. Copyright 2020, Reproduced with permission from American Institute of Physics Publishing.

polymer chains, polymer crystal structures, and polymer amorphous structures. Adoption of this tool has accelerated the development of datasets for gas transport in polymers [87], and may be used in the future to generate data relevant to high-temperature dielectrics.

5.2. Polymer representations

After compiling the data set, the next step is to represent the key chemical aspects of each polymer in a machine-ingestible manner. Modern approaches to representing (i.e., fingerprinting) polymers rely on the SMILES language [101], which is used to represent 2D chemical structures as text. The string of text corresponding to one structure is a so-called "SMILES string". Another key development has been the RDKit package [102], which, via a SMILES string, can generate a 2D mathematical graph, with atoms as nodes and bonds as edges, for any given molecule. The representation of a 2D chemical structure as a graph allows for the calculation of certain chemical features to be automated and completed efficiently. One such group of features is contained in the Extended-Connectivity Fingerprint (ECFP) [91,103]. Initially developed to address the mathematical challenge of molecular isomorphism (the task of identifying instances where two molecules, with different atom numberings, are the same), ECFP accurately identifies unique chemical structures. Subsequently, ECFP has been used as a fingerprint for training polymer structure-property models. However, the ECFP, when applied to graphs of polymer repeat units, is not invariant to relevant transformations such as translation, addition, and subtraction. The process of translation involves shifting the periodicity window, leading to equivalent periodic repeat units, such as (−OCC−), (−COC−), and (−CCO−) in polyethylene glycol. Addition refers to extending a repeat unit by incorporating one or more minimal repeat units, exemplified by (−COCO−) and (−COCOCO−). Subtraction is the opposite, where a repeat unit is reduced by removing minimal repeat units.

Over the past decade, a successful fingerprinting approach has been developed as part of the Polymer Genome (PG) project [92]. Importantly, the fingerprint generated by this approach is invariant to translation, addition, and subtraction. The input to the fingerprint algorithm is the SMILES string corresponding to a polymer's repeat unit. The algorithm is characterized by a key feature: hierarchical representation, organized across physical length scales (see Fig. 13). At the smallest length scale, the atomic features of a polymer are considered. Specifically, we check the frequency of various small substructures, known as atomic triples. Moving up a rung, the block-level features of a polymer are considered. This corresponds to checking the frequency of large substructures within a polymer. Finally, at the largest length scale, chain-level features—spanning the entire polymer repeat unit—are computed. Some example chain-level features include the distance between rings, side-chain length, and the number of aromatic rings. The PG approach works well because it encapsulates chemical features supported by decades of science and natural intelligence. For example, side-chain length, one of the features in the PG fingerprint, is known to impact several properties, including glass-transition temperature.

5.3. Structure-Property prediction models

Once the data set has been compiled and fingerprinted, a structure-property prediction model may be trained. During the training process, a set of relationships is learned between the fingerprint components and the properties are interest, such that the relationships can best fit the fingerprinted data set. This set of learned relationships *is* the structure-property model and, once obtained, may be used to predict the properties of known and unknown structures.

Our initial work [104] in this area focused on the prediction of five properties (atomization energy E_{at}, band gap E_g, and the electronic, ionic, and total dielectric constants) using the kernel ridge regression (KRR) machine learning algorithm and the atomic-level components of the PG fingerprint. In the intervening time, additional layers of sophistication have been added to model the aforementioned five properties and more (e.g., T_g, electron injection barrier, etc.). The advanced due to the curation of larger data sets, [93–96,105] advancements beyond KRR (Gaussian Process Regression, [106–108] neural networks, [93,95,96,98], etc.), multi-task learning, [96] and improved fingerprinting techniques [93,94] (many of the advanced models are currently deployed at https://polymergenome.org/).

While KRR is a useful method, it is only able to learn linear relationships between fingerprints and properties. GPR, on the other hand, can pick up nonlinear relationships and was employed in several works for this reason. The accuracy of models trained using either method suffers when data is scarce or of poor quality. To mitigate this issue, we have utilized multitask learning, wherein instead of training one model per property, a singular model is trained to learn multiple properties simultaneously. By training a

a PolyGNN – A graph neural network for polymers
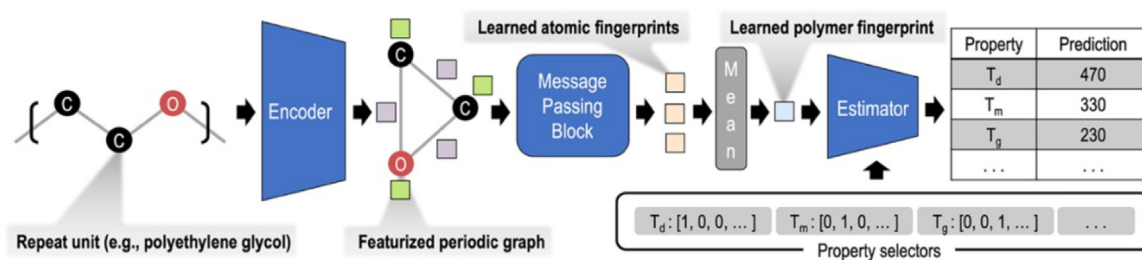
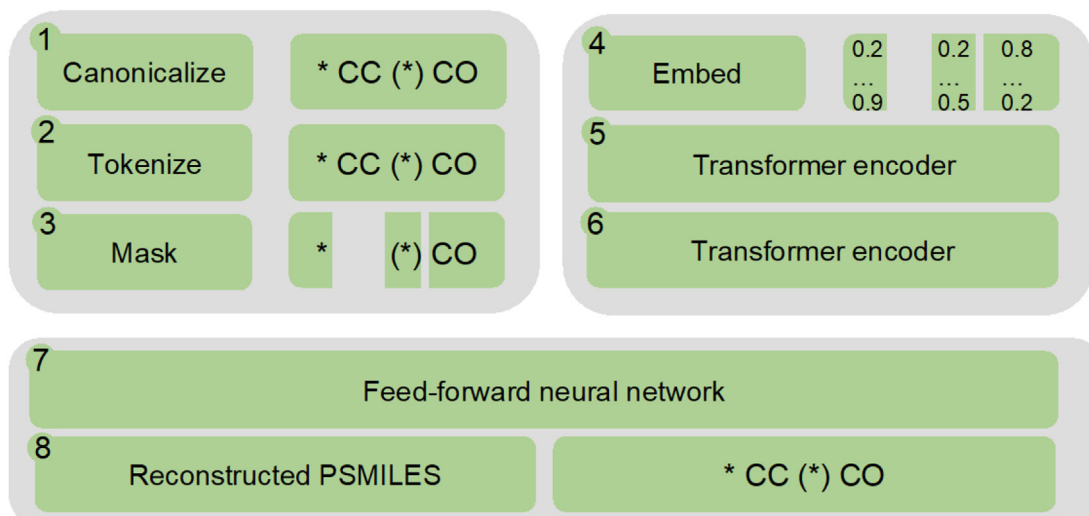## b PolyBERT – A Transformer model for polymers

Fig. 14. a A graph neural network (termed polyGNN) for polymers. The starting point is the repeat unit, which is then passed into the polyGNN encoder. The encoder generates initial features for each atom and bond in the graph. These atom and bond features are passed to the polyGNN Message Passing Block, which updates the atom features using a set of MLPs. The updated atom features are averaged together, yielding the polymer features. These features are input, along with selector vectors, into the polyGNN Estimator (another MLP) which yields the property prediction [93]. Copyright 2023, Reproduced with permission from American Chemical Society. b polyBERT is designed to process Polymer Simplified Molecular-Input Line-Entry System (PSMILES) strings. It performs several key steps, including canonicalization, tokenization, and masking (steps 1–3) of PSMILES strings, before passing them into the DeBERTa [113] model for further processing (steps 4–6). The architecture includes 12 Transformer encoders, each equipped with 12 attention heads. To predict masked tokens, a final dense layer with a softmax activation function is employed. Additionally, polyBERT generates fingerprints by averaging over the token dimension (sentence average) of the last Transformer encoder, to create meaningful representations for downstream applications [94]. Copyright 2023, Adapted with permission from Springer Nature.

single model to simultaneously learn a group of target properties at once, the risk of generating overfitted predictions for any one specific target property is reduced [109,110]. As a result, the accuracy of each property is improved. For instance, for 24 out of 34 polymer properties (~71 %), models leveraging multitask learning outperformed their single task counterparts [93].

An important aspect of multitask involves carefully ensuring that all properties selected for co-learning are interrelated [96]. This selection is a heuristic process that relies on materials domain expertise. Failing to establish such links may result in learning spurious correlations, thereby impairing the model's ability to generalize effectively. A challenge with multitask learning is that the number of training data points n can be quite large. In these cases, GPR is an inefficient choice, as it requires a $n \times n$ matrix inversion during training. Neural networks offer improved efficiency.

Over the past few decades, huge amounts of research and investment have gone into producing ever-more-performant neural network architectures. The most basic neural network is a multi-layer perceptron (MLP), which takes as input a fixed-length vector (e.g., the PG fingerprint, ECFP, etc.). More advanced architectures include the graph neural network (GNN) for processing graph data [111] and the Transformer [112] for processing text. As mentioned earlier, polymers may be represented using either data type and therefore both the GNN and the Transformer architectures may be used to produce polymer structure-property models. In these models, unlike with the MLP, there is no need to calculate the fingerprint components before training, reducing the computational overhead. In our work (Fig. 14), we have found that both GNN and Transformer-based polymer prediction algorithms are up to two orders of magnitude faster than MLP-based approaches with comparable or better model accuracy [93,94]. The enhanced computational efficiency of GNN and Transformer architectures is expected to drive their broader adoption, unlocking the ability to screen extremely large polymer libraries at scale.
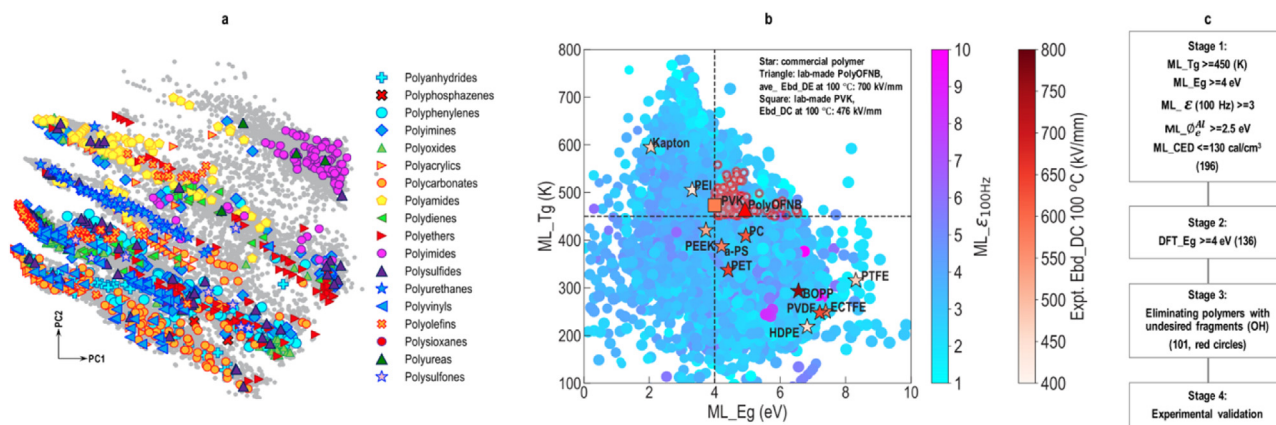
Fig. 15. a Chemical space of 13,000 previously synthesized polymers, displayed with the help of the first two principal components (PC1 and PC2). b Property heat map of machine learning glass transition temperature, band gap, dielectric constant, and electron injection barrier for aluminum electrode - ML Tg, ML Eg, ML ε100 Hz, and ML $\varphi_e{}^{Al}$ (circles sizes) respectively of 13,000 polymers. (The experimental E_{bd} DC values at 100 °C of 11 commercial polymers (stars) and lab-made PVK (square) and the average D−E loop Ebd of PolyOFNB (triangle) are shown). c Down-selection procedure for polymer dielectrics screening [114]. Copyright 2021, Reproduced with permission from the American Chemical Society.
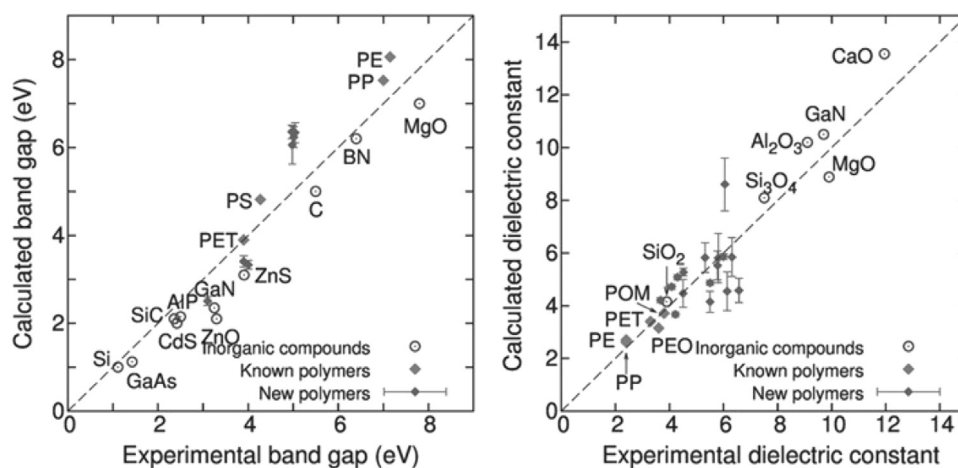
Fig. 16. The calculated bandgaps and dielectric constants from density functional theory (DFT) computations are compared against experimental measurements for selected inorganic compounds and polymers [42]. Copyright 2016, Reproduced with permission from John Wiley & Sons Inc.

5.4. Structure generation

The space of known polymers is one such library. In our recent work, [114] starting from a list of 13,000 known polymers, we used an informatics-based approach to identify promising dielectrics. The space of polymers in our data set, displayed in Fig. 15a, spans over 18 polymer families, including some that have not previously been tested for dielectric properties. Promising dielectrics should exhibit high T_g and high E_g. While these properties are uncommon for commercial polymer dielectrics (stars in Fig. 15b), our structure-property models reveal interesting candidates (dots in Fig. 15b) that are known polymers but have not yet been commercialized as dielectrics.

To focus on the most promising of these candidates, a four-stage approach is used, detailed in Fig. 15c. In the first stage, after models are trained for five properties (glass transition temperature, band gap, dielectric constant at 100 Hz, electron injection barrier for aluminum electrode, and cohesive energy density - T_g, E_g, $\varepsilon_{100\ Hz}$, φ_eAl, and CED), we keep the 196 structures with predicted Tg $\geq$ 450 K, $E \geq$ 4 eV, $\varepsilon_{100\ Hz} \geq$ 3, φ_eAl $\geq$ 2.5 eV, and CED $\leq$ 130cal/cm3. For these 196 structures, we compute the Eg using DFT, keeping the 136 structures with computed Eg $\geq$ 4 eV. In the third stage, we eliminate structures with undesirable, as determined by polymer chemists, functional groups (e.g., OH), leav-

ing 101 candidates. Finally, in the fourth stage, a handful of candidates are synthesized and characterized in the lab. Among these, we identified polyvinylcarbazole (PVK) and discovered that it has a relatively high Ebd of 476 kV/mm at 100 °C, thus making it a promising candidate for applications requiring high temperature, high energy density capacitors.

The aforementioned 13,000 known polymers represent a speck in the universe of possible polymers. Several methods have been developed and used to chart these unknown regions. A polymer repeat unit may be viewed as a sequence of connected chemical fragments (e.g., NH, CO, CS, etc.). By identifying suitable fragments and combining them, a list of candidate polymers can be enumerated. Following this approach, we identified polymers with exceptional predicted band gap and dielectric constant (see "New polymers" in Fig. 16) [42]. These "first-generation" structures were used as starting points to identify promising polymer families. We synthesized a set of "second-generation" polymers within these families, and found a few (see Fig. 11) with unprecedented energy densities, up to 16 J/cm^3 at room temperature.

Rather than generating polymers from all possible fragment combinations, another approach is to use an optimization algorithm to navigate and selectively screen the space. One optimization algorithm that we have used is the genetic algorithm (GA) [97,98], described, for polymers, in Fig. 17a. In this evolution-
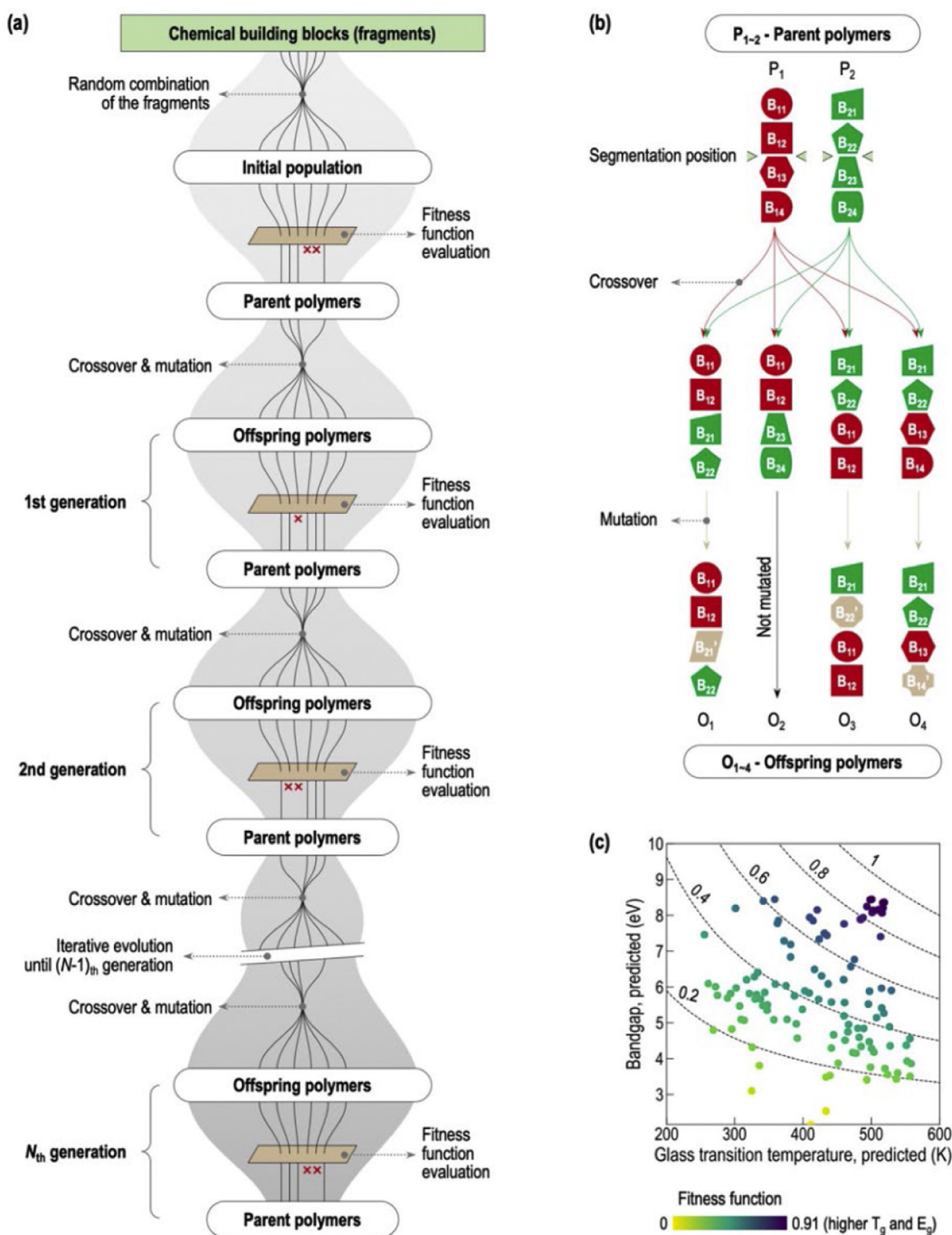
Fig. 17. (a) Illustrates the iterative workflow for polymer design using a genetic algorithm framework. (b) Depicts the mutation and crossover operations to generate new 'offspring' polymer structures from a pair of 'parent' polymers. (c) Shows a plot of glass transition temperature (Tg) versus bandgap (Eg) for the offspring polymer population [98]. Copyright 2021, Reproduced with permission from Elsevier Science Ltd.

inspired scheme, the starting point is a set of chemical fragments, analogous to the set of base pairs found in DNA. An initial population of "parent" polymers is created by a random combination of the fragments. A fitness function is used to select the most fit parents. Crossover and mutation (see Fig. 17b) of the parents' chemical fragments produce "offspring polymers". These offspring serve as parents for the next generation. This iterative process continues until the goal (high-performance dielectrics in this case) has been met. In this work, structure-property models are leveraged to assess fitness in each iteration of the optimization process.

This GA has been used to produce polymers with high predicted E_g, high predicted T_g, and beyond (see Figs. 17c and 18). Of particular interest are the polymers listed in Fig. 18, which were screened based on both predicted properties and synthesizability. Each polymer is predicted to display an $E_g > 5$ eV, a dielectric constant at 100 Hz $\epsilon > 4$, a $T_g > 500$ K, a cohesive energy density $e_{coh} \langle$ 70 cal cm^{-3}, and an electron injection barrier with Al electrode $\phi_e \rangle 3$ eV. In addition, Fig. 18 lists potential monomers and a synthesizability score for each polymer. The monomers and scores were generated using a retrosynthesis tool developed by us [115]. Using the tool, out of the large number of generated polymers with good
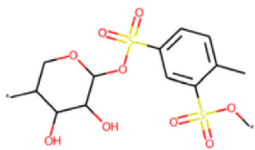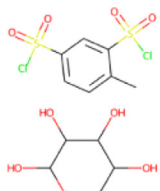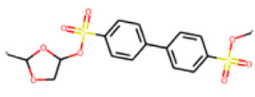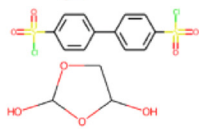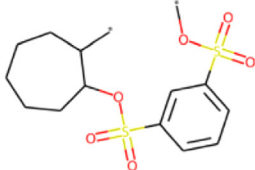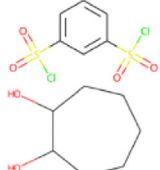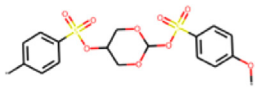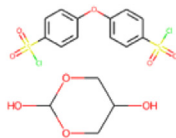
Polymer	Structure	Properties	Predicted reactants	S_{score}
1		E_g: 5.6 ± 0.4 eV ε: 4.7 ± 1.5 T_g: 526 ± 91 K e_{coh}: 47 ± 90 cal cm^{-3} Φ_e: 3.2 ± 0.2 eV		0.88
2		E_g: 5.1 ± 0.5 eV ε: 4.2 ± 0.4 T_g: 539 ± 88 K e_{coh}: 57 ± 85 cal cm^{-3} Φ_e: 3.04 ± 0.17 eV		0.80
3		E_g: 5.3 ± 0.4 eV ε: 4.5 ± 0.5 T_g: 533 ± 86 K e_{coh}: 51 ± 88 cal cm^{-3} Φ_e: 3.1 ± 0.19 eV		0.77
4		E_g: 5.2 ± 0.5 eV ε: 4.3 ± 0.4 T_g: 530 ± 81 K e_{coh}: 67 ± 81 cal cm^{-3} Φ_e: 3.2 ± 0.16 eV		0.71

Fig. 18. Polymers derived using a genetic algorithm for achieving the optimal values for intended proxies [97]. Copyright 2021, Reproduced with permission from Springer Nature.

predicted properties, the ones with available monomers were efficiently identified.

Another optimization algorithm useful for polymer design is gradient descent, which is used to train generative neural networks capable of design. A benefit of these approaches is that, in principle, structure-property models are not required. In practice, we find that incorporating structure-property models into the design process improves performance, although the models need not be used iteratively (as was required with the GA). The first approach we have explored is the Syntax-Directed Variational Autoencoder (SD-VAE) [99,116]. A VAE is a neural network that learns a mapping between a discrete input sequence (i.e., a polymer SMILES string in this case) and a continuous "latent" space. The VAE consists of an Encoder, to learn the input to latent mapping, and a Decoder, to learn the latent to input mapping. Thus, by sampling the latent space and then using the Decoder, SMILES strings corresponding to new polymers may be generated. Notably, the VAE is an unsupervised learning method that only requires polymer SMILES strings for training, without any associated property labels.

SMILES strings have syntactical and semantic rules that must be followed for the string to represent a polymer. These rules are incorporated into the VAE using grammar parse trees and stochastic lazy attributes, details of which can be found in Ref. [99]. To locate new polymers with high E_g and T_g, we sampled points in the latent space that lie on straight lines between the latent space vectors of known polymers with high E_g and T_g. The samples were screened using E_g and T_g structure-property models. A handful of the screened candidates are displayed in Fig. 19.

Another generative neural network-based approach for polymer design has been developed by us. The algorithm is called Polymer Graph-to-Graph translation (polyG2G) [100]. In the SD-VAE approach, the Decoder was trained to reproduce the sequence fed into the decoder. Here, the Decoder is trained to produce not the same sequence, but a closely-related sequence. Specifically, the model is trained to translate the input sequence, which represents a polymer with low dielectric performance, into a chemically similar polymer with much better dielectric performance.

The training methodology for the polyG2G system involves several intricate steps, as illustrated in Fig. 20 (steps A-J). At a high level, the polyG2G process starts by mapping a polymer to both a junction tree (JT) and graph representation. These representations are encoded into latent vectors, concatenated, and sampled to produce translated Jts and graphs. These are then decoded into new polymer candidates by determining the constituent atoms, rings, and chemical bonds. During training, the objective is to generate a graph sufficiently similar to the target graph, with the training loss measuring the success of this objective. During inference, the trained model performs the same steps but focuses on generating new polymers.

Once trained, the polyG2G model was used to design polymers with $E_g > 4$ eV, $\phi_e > 3$ eV, and $T_g > 450$ K. Our analysis identified 3556 unique polymers (0.45 % of the total) that met our predefined objectives. In contrast, only 8 out of 13,014 (0.061 % of the total) polymers from the dataset used to train polyG2G satisfied our objectives. This stark difference demonstrates that polyG2G is not only capable of generating a large number of high-performing designs but also that these designs "hit" the target objectives an order of magnitude more frequently than the space of synthesized polymers. In other words, polyG2G can effectively learn targeted design rules specifically aimed at producing high dielectric breakdown strength polymers.

To further refine our search, we subjected 20 of the 3556 polymers to density functional theory (DFT) computations to calculate their bandgap and electron injection barrier. These 20 candidates were specifically chosen from the larger pool as they exhibited the
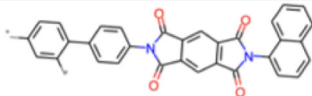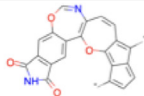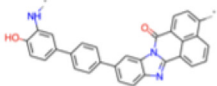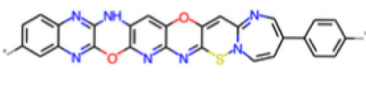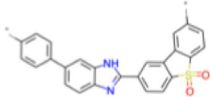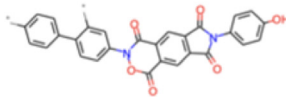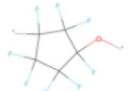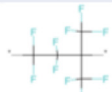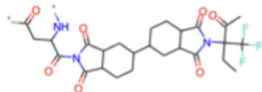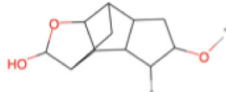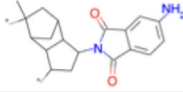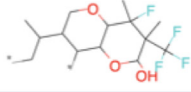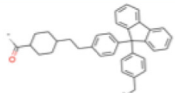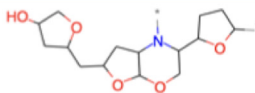
#	Example New Polymers	
	Enumeration	Generation
High Tg (in K)		
1		
	686.77(53.78)	737.76 (57.30)
2		
	661.90(54.18)	708.59(70.13)
3		
	649.04(52.20)	707.83(53.64)
High Eg (in eV)		
1		
	8.23 (1.04)	8.61 (1.10)
2		
	7.96 (1.02)	8.56 (1.06)
3		
	7.60 (1.02)	8.46 (1.01)
High Tg (in K), Eg (in eV)		
1		
	517.44 (81.82), 4.97 (1.67)	509.53 (54.01), 6.32 (1.11)
2		
	501.54 (65.21), 5.49 (1.34)	528.65(56.11), 6.05 (1.19)
3		
	548.52 (56.93), 4.61 (1.30)	506.89 (52.82), 6.2 (1.11)

Fig. 19. Polymers developed using the SD-VAE approach and their Gaussian process regression (GPR) Tg and Eg estimates and error (in parenthesis). ('*' is a representative of polymer chain ends) [99]. Copyright 2020, Reproduced with permission from the American Chemical Society.
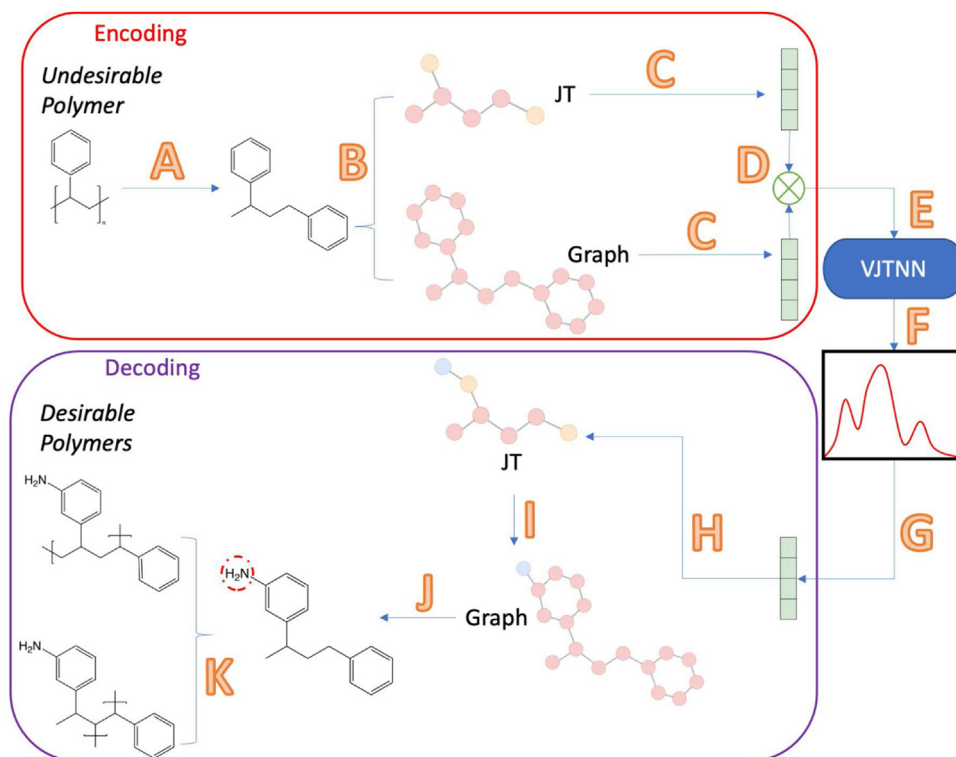
Fig. 20. The polyG2G workflow for the example case of polystyrene, where the hyperparameters npair and ntranslate are set to 2 and 1, respectively. The result of the latent translation step is encircled by a dashed red line. The pink circles represent carbon atoms, the yellow circles denote benzene rings, and the blue circles symbolize NH2 groups. All hydrogen atoms are implicit in this representation [100]. Copyright 2021, Reproduced with permission from the American Chemical Society.

highest fitness score F, as defined in Equation 4.

$$F = T_g \times E_g \times \varphi_e \times R_{mt} \times U \qquad (3)$$

Where Tg, E_g, ϕ_e are the ML property predictions of a given polymer, R_{mt} is the ratio between the number of atoms in the main chain per repeat unit and the total number of atoms per repeat unit, and U is the uniqueness of the polymer.

After subjecting these 20 candidates to DFT computations, 10 polymers exhibited computed properties that surpassed our predefined objectives, indicating their resistance to large electric fields. These top-performing candidates are presented in Table 3.

6. Critical next steps

6.1. Beyond homopolymers

Homopolymers possess structural simplicity; they can be described by a single repeating unit. However, neat copolymers and polymer blends, which are structurally stochastic, are much more common in everyday life. Polymer composites, in which the polymer (homo-, co-, and/or blend) matrix forms the primary continuous phase, are even more common. In the context of dielectrics, each polymeric material class has its merit [117]. However, to date, most work in polymer informatics has focused on homopolymers. Limited work has been done on copolymers [118–124], blends [109,125–127], and polymer composites [128]. These efforts have elucidated techniques for representing complex polymer systems. These techniques now need to be applied to model properties relevant to dielectric design. For example, while models [95] exist for the glass-transition temperature of complex polymer systems, there are not yet curated data sets or models for the dielectric constant and band gap of complex polymer systems.

6.2. Synthesis and processing

As detailed in prior sections, a handful of methods have been developed for designing or enumerating novel polymer repeat units. In these methods, the repeat units are produced atom-by-atom or by combining molecular fragments. However, ranking such repeat units in terms of their synthesizability remains a highly nontrivial task. For example, a polymer with repeating unit $-N_{R1}$-$N_{R2}-$ having a nitrogen backbone with R_1, and R_2, being substituent functional groups might have a high Tg, and dielectric constant theoretically, but the synthesis of such polymer is challenging, especially on an industrial scale.
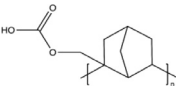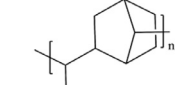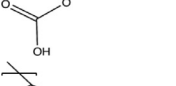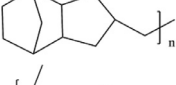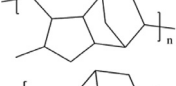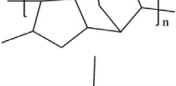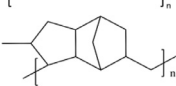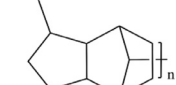
A promising alternative approach [75,129,130] is to start with a library of synthesizable small molecules, and then digitally transform them into repeat units using a set of well-known polymerization reactions. This way, some aspects (e.g., the monomers and the reaction type) of the polymer synthesis are known *a priori* for each repeat unit.

Computational assessment of polymer synthesis or processing parameters will also be critical. To date, in the context of polymer dielectrics, computations have mainly been used to predict the properties (e.g., Tg, band gap, dielectric constant) of polymers in the solid state (e.g., dielectric films). Limited work exists on computations for dielectric polymers in other states (e.g., in solution or melt), however, there is a tremendous opportunity to accelerate dielectric design by looking beyond the solid state.

For instance, prior work has focused on predicting the compatibility of polymers in a variety of solvents [131–134]. These models may be extended to future studies of dielectric polymers. The models can accelerate the choice of solvent (and minimize solvent waste) for a given polymer dielectric during chemical synthesis or even help design new polymers that are compatible with targeted (e.g., green) solvents.

Table 3
Structures and properties of 10 novel, high-value, targets discovered by polyG2G. Band gaps (E_g) and electron injection barriers (ϕ_e) are DFT estimates. Glass-transition temperatures (T_g) are ML estimates [100]. Copyright 2021, Reproduced with permission from the American Chemical Society.

Structure	Properties		
	E_g, DFT (eV)	ϕ_e, DFT (eV)	T_g, ML (K)
	6.04	3.13	549
	6.52	3.05	520
	6.21	3.27	480
	6.07	3.12	511
	6.14	3.08	495
	6.28	3.24	502
	6.32	3.24	474
	6.07	3.08	524
	6.26	3.19	478
	6.29	3.17	499

Compared to polymers in solution, the body of informatics work on polymers in melt is limited and has mainly focused on modeling the polymer melting temperature [93,135]. Other models exist for properties such as melt quality [136] and melt instability, [137] but the models were trained on just a handful of polymer chemistries and therefore do not generalize to new polymeric materials. Future work focused on creating general-purpose models for polymer melt properties is needed.

6.3. Data collection

Many of the critical next steps identified above depend on the presence of yet-to-be-developed data sets. There are two pathways. The first is to curate data that already exists, albeit among several disparate sources, into one place. Advancements (e.g., the Transformer [112], ChatGPT [138], BERT [139]) in the field of Natural Language Processing have led to the creation of Large Language Models (LLMs) that can reliably perform complex natural language tasks. These tasks include sentiment analysis [140], machine trans-

lation [141], and spam detection [142]. LLMs have also been used to mine literature data on polymeric materials from published abstracts [143]. Expansion of this capability to more abstracts, full-text, tables, and graphs is a promising direction for obtaining necessary, dielectric-relevant, data.

The other data-collection pathway is generating new data, which may be done using high-throughput simulations (e.g., MD or DFT). Such data is unlikely to replace ground-truth experimental data in terms of fidelity. However, using the multi-task learning schemes described in previous sections, the simulation data may be used as a complement to improve models that are only trained on experimental data.

6.4. Physics-informed machine learning

While low barriers exist to curating and generating large data sets for some properties (e.g., by NLP or simulation), there are other properties for which data is scarce. In these scenarios, physical or empirical equations may be used to decrease an ML model's dependency on data. For instance, an Arrhenius relationship was incorporated into a deep neural network model for the prediction of lithium-ion conductivity in polymer electrolytes [144]. By incorporating physics (in the form of the Arrhenius relationship), the model produced smooth and meaningful predictions of conductivity as a function of temperature, despite only a limited amount of temperature variation in the training data.

Likewise, there exist well-known relationships for important dielectric properties, such as the Fox Equation, [145] for glass-transition temperature. Future work that incorporates the time-tested wisdom of these equations will improve the generalizability and reliability—especially in low-data situations—of next-generation ML models for dielectric design.

7. Challenges and future directions

Structural modifications for improving the capacitive energy storage can simultaneously influence a variety of other polymer properties. While it is important to improve the thermal stability (Tg) to avoid thermal degradation, increasing Tg can introduce challenges in polymer processing, such as increased film brittleness and internal stress [3]. Similarly, enhanced dielectric polarization to improve the energy density may inadvertently elevate dielectric loss, which can degrade performance efficiency over multiple operation cycles [146–148]. Excessive dielectric loss is particularly concerning for large-scale applications of capacitor films, where even a small increase in loss can result in significant heating of the electrical assembly due to the large surface area of the film. The introduction of polar groups for enhancing the dielectric constant can decrease the moisture resistance of the polymer film if the functional groups are hydrophilic. The presence of unwanted moisture can lead to an unstable dielectric constant, which is not desired for capacitive energy storage leading to early dielectric breakdown. Several such structural modifications can also affect the polymer-solvent interactions, film morphology, mechanical properties, etc. Also, the mechanism of dielectric breakdown and the direct effect of thermal conductivity on capacitive energy storage performance is not fully understood [149].

Properties such as Tg, dielectric loss, mechanical strength, and thermal and electrical breakdown are inherently complex and multi-variable, depending on the polymer morphology, multi-scale geometric architectures, defects of different sizes, etc. AI-based tools that offer economical screening of existing polymers, and build on to discover new polymers can be instrumental in understanding these properties across multiple variables [150]. Well-trained and calibrated artificial intelligence (AI) models, can be used to compute the calculation of large numbers that are difficult

to manually comprehend, and unravel the mutual interdependence of various structure-property relations.

8. Conclusion

The co-design approach has the potential to accelerate the research and develop technologies that can moderate the rising energy-storage demand. Machine learning techniques can significantly reduce the time required to develop new polymers and to elucidate the interdependencies of various parameters following structural modifications. For instance, through initiatives like the Material Genome Initiative (MGI), the time required to develop new materials could be cut by more than half compared to conventional methods [151].

In conclusion, it is important to highlight the promising potential of a systematic method for the discovery of new polymer dielectrics. Recent achievements in the field of computer engineering and material science demonstrate that innovative collaborative methods integrating computation, synthesis, and characterization can lead to significant progress [152–157]. The co-design approach enables us to move ahead of the traditional trial-and-error and instinct-based methods that have been used so far.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Pritish S Aklujkar: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Rishi Gurnani:** Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Methodology, Investigation, Formal analysis, Data curation. **Pragati Rout:** Writing – review & editing, Visualization. **Ashish R Khomane:** Writing – review & editing. **Irene Mutegi:** Writing – review & editing. **Mohak Desai:** Writing – review & editing. **Amy Pollock:** Writing – review & editing. **John M Toribio:** Writing – review & editing. **Jing Hao:** Validation. **Yang Cao:** Writing – review & editing, Visualization, Validation, Supervision, Project administration, Funding acquisition, Conceptualization. **Rampi Ramprasad:** Writing – review & editing, Visualization, Validation, Supervision, Project administration, Methodology, Conceptualization. **Gregory Sotzing:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Data availability

Data will be made available on request.

References

[1] Sotzing GA, Aklujkar PS. Chromogenic identification of breakdown. Nat Mater 2024;23:163–4. doi:10.1038/s41563-023-01786-9.
[2] Liu C, Li F, Ma LP, Cheng HM. Advanced Materials for energy storage. Adv Mater 2010;22:E28–62. doi:10.1002/adma.200903328.
[3] Feng QK, Zhong SL, Pei JY, Zhao Y, Zhang DL, Liu DF, et al. Recent progress and future prospects on all-organic polymer dielectrics for energy storage capacitors. Chem Rev 2022;122:3820–78. doi:10.1021/acs.chemrev.1c00793.
[4] Yang M, Ren W, Guo M, Shen Y. High-energy-density and high efficiency polymer dielectrics for High temperature electrostatic energy storage: a review. Small 2022;18. doi:10.1002/smll.202205247.
[5] Chen Q, Shen Y, Zhang S, Zhang QM. Polymer-based dielectrics with high energy storage density. Annu Rev Mater Res 2015;45:433–58. doi:10.1146/annurev-matsci-070214-021017.
[6] Pei J, Yin L, Zhong S, Dang Z. Suppressing the loss of polymer-based dielectrics for high power energy storage. Adv Mater 2023;35:2203623. doi:10.1002/adma.202203623.
[7] Yang Z, Yue D, Yao Y, Li J, Chi Q, Chen Q, et al. Energy storage application of all-organic polymer dielectrics: a review. Polymers (Basel) 2022;14:1160. doi:10.3390/polym14061160.
[8] Ahmad Z. Polymer dielectric materials. Dielectric materials. Silaghi MA, editor. InTechOpen; 2012. doi:10.5772/50638.
[9] Tan D, Zhang L, Chen Q, Irwin P. High-temperature capacitor polymer films. J Electron Mater 2014;43:4569–75. doi:10.1007/s11664-014-3440-7.
[10] Chung TM. Functionalization of polypropylene with high dielectric properties: applications in electric energy storage. Green Sustain Chem 2012;02:29–37. doi:10.4236/gsc.2012.22006.
[11] Wu C, Deshmukh AA, Li Z, Chen L, Alamri A, Wang Y, et al. Flexible temperature-invariant polymer dielectrics with large bandgap. Adv Mater 2020;32. doi:10.1002/adma.202000499.
[12] Li Q, Zhang G, Liu F, Han K, Gadinski MR, Xiong C, et al. Solution-processed ferroelectric terpolymer nanocomposites with high breakdown strength and energy density utilizing boron nitride nanosheets. Energy Environ Sci 2015;8:922–31. doi:10.1039/C4EE02962C.
[13] Li Q, Chen L, Gadinski MR, Zhang S, Zhang G, Li HU, et al. Flexible high-temperature dielectric materials from polymer nanocomposites. Nature 2015;523:576–9. doi:10.1038/nature14647.
[14] Wu C, Deshmukh AA, Chen L, Ramprasad R, Sotzing GA, Cao Y. Rational design of all-organic flexible high-temperature polymer dielectrics. Matter 2022;5:2615–23. doi:10.1016/j.matt.2022.06.064.
[15] Yue D, Yin J, Zhang W, Cheng X, Zhang M, Wang J, et al. Computational simulation for breakdown and energy storage performances with optimization in polymer dielectrics. Adv Funct Mater 2023;33. doi:10.1002/adfm.202300658.
[16] Zhu M, Deng T, Dong L, Chen J, Dang Z. Review of machine learning-driven design of polymer-based dielectrics. IET Nanodielectr 2022;5:24–38. doi:10.1049/nde2.12029.
[17] Odian G. Principles of polymerization. Wiley; 2004. doi:101002/047147875X.
[18] Gedde UW. Polymer physics. Dordrecht: Springer Netherlands; 1999. doi:101007/978-94-011-0543-9.
[19] Baldwin AF, Ma R, Wang C, Ramprasad R, Sotzing GA. Structure-property relationship of polyimides based on pyromellitic dianhydride and short-chain aliphatic diamines for dielectric material applications. J Appl Polym Sci 2013;130:1276–80. doi:10.1002/app.39240.
[20] Shen Y, Nan CW. High thermal conductivity dielectric polymers show record high capacitive performance at high temperatures. Natl Sci Rev 2023;10. doi:10.1093/nsr/nwad224.
[21] Chen J, Zhou Y, Huang X, Yu C, Han D, Wang A, et al. Ladderphane copolymers for high-temperature capacitive energy storage. Nature 2023;615:62–6. doi:10.1038/s41586-022-05671-4.
[22] Li Q, Liu F, Yang T, Gadinski MR, Zhang G, Chen LQ, et al. Sandwich-structured polymer nanocomposites with high energy density and great charge–discharge efficiency at elevated temperatures. Proc Natl Acad Sci 2016;113:9995–10000. doi:10.1073/pnas.1603792113.
[23] Fan B, Zhou M, Zhang C, He D, Bai J. Polymer-based materials for achieving high energy density film capacitors. Prog Polym Sci 2019;97:101143. doi:10.1016/j.progpolymsci.2019.06.003.
[24] Ho JS, Greenbaum SG. Polymer capacitor dielectrics for high temperature applications. ACS Appl Mater Interfaces 2018;10:29189–218. doi:10.1021/acsami.8b07705.
[25] Luo B, Wang X, Wang Y, Li L. Fabrication, characterization, properties and theoretical analysis of ceramic/PVDF composite flexible films with high dielectric constant and low dielectric loss. J Mater Chem A 2014;2:510–19. doi:10.1039/C3TA14107A.
[26] Kim C, Pilania G, Ramprasad R. From organized high-throughput data to phenomenological theory using machine learning: the example of dielectric breakdown. Chem Mater 2016;28:1304–11. doi:10.1021/acs.chemmater.5b04109.
[27] Zakrevskii VA, Sudar NT. Ionization mechanism of the electrical degradation (breakdown) of polymer dielectric films. Phys Solid State 2013;55:1395–400. doi:10.1134/S1063783413070354.
[28] Sun Y, Bealing C, Boggs S, Ramprasad R. 50+ years of intrinsic breakdown. IEEE Electr Insulat Mag 2013;29:8–15. doi:10.1109/MEI.2013.6457595.
[29] Grabowski CA, Fillery SP, Koerner H, Tchoul M, Drummy L, Beier CW, et al. Dielectric performance of high permitivity nanocomposites: impact of polystyrene grafting on BaTiO 3 and TiO 2. Nanocomposites 2016;2:117–24. doi:10.1080/20550324.2016.1223913.
[30] Zhang B, Liu J, Ren M, Wu C, Moran TJ, Zeng S, et al. Reviving the "Schottky" barrier for flexible polymer dielectrics with a superior 2D nanoassembly coating. Adv Mater 2021;33:2101374. doi:10.1002/adma.202101374.
[31] Kaltenbrunner M, Sekitani T, Reeder J, Yokota T, Kuribara K, Tokuhara T, et al. An ultra-lightweight design for imperceptible plastic electronics. Nature 2013;499:458–63. doi:10.1038/nature12314.
[32] Forrest SR. The path to ubiquitous and low-cost organic electronic appliances on plastic. Nature 2004;428:911–18. doi:10.1038/nature02498.

[33] Artbauer J. Electric strength of polymers. J Phys D Appl Phys 1996;29:446–56. doi:10.1088/0022-3727/29/2/024.

[34] Kamal D, Wang Y, Tran HD, Chen L, Li Z, Wu C, et al. Computable bulk and interfacial electronic structure features as proxies for dielectric breakdown of polymers. ACS Appl Mater Interfaces 2020;12:37182–7. doi:10.1021/acsami.0c09555.

[35] Li H, Zhou Y, Liu Y, Li L, Liu Y, Wang Q. Dielectric polymers for high-temperature capacitive energy storage. Chem Soc Rev 2021;50:6369–400. doi:10.1039/D0CS00765J.

[36] Li Z, Xu C, Uehara H, Boggs S, Cao Y. Transient characterization of extreme field conduction in dielectrics. AIP Adv 2016;6. doi:10.1063/1.4971158.

[37] Ieda M. Electrical conduction and carrier traps in polymeric materials. IEEE Trans Electr Insulat 1984;EI-19:162–78. doi:10.1109/TEI.1984.298741.

[38] Chen L, Batra R, Ranganathan R, Sotzing G, Cao Y, Ramprasad R. Electronic structure of polymer dielectrics: the role of chemical and morphological complexity. Chem Mater 2018;30:7699–706. doi:10.1021/acs.chemmater.8b02997.

[39] Ho J, Jow TR. High field conduction in biaxially oriented polypropylene at elevated temperature. IEEE Trans Dielectr Electr Insulat 2012;19:990–5. doi:10.1109/TDEI.2012.6215104.

[40] Alamri A, Wu C, Mishra A, Chen L, Li Z, Deshmukh A, et al. Improving the rotational freedom of polyetherimide: enhancement of the dielectric properties of a commodity high-temperature polymer using a structural defect. Chem Mater 2022;34:6553–8. doi:10.1021/acs.chemmater.2c01441.

[41] Wu C, Alamri A, Deshmukh AA, Li Z, Islam S, Sotzing GA, et al. A modified polyetherimide film exhibiting greatly suppressed conduction for high-temperature dielectric energy storage. IEEE Conf Electr Insulat Dielectr Phenom 2020;1–4. doi:10.1109/CEIDP49254.2020.9437471.

[42] Mannodi-Kanakkithodi A, Treich GM, Huan TD, Ma R, Tefferi M, Cao Y, et al. Rational Co-design of polymer dielectrics for energy storage. Adv Mater 2016;28:6277–91. doi:10.1002/adma.201600377.

[43] Huan TD, Boggs S, Teyssedre G, Laurent C, Cakmak M, Kumar S, et al. Advanced polymeric dielectrics for high energy density applications. Prog Mater Sci 2016;83:236–69. doi:10.1016/j.pmatsci.2016.05.001.

[44] Lorenzini RG, Kline WM, Wang CC, Ramprasad R, Sotzing GA. The rational design of polyurea & polyurethane dielectric materials. Polymer 2013;54:3529–33. doi:10.1016/j.polymer.2013.05.003.

[45] Ma R, Baldwin AF, Wang C, Offenbach I, Cakmak M, Ramprasad R, et al. Rationally designed polyimides for high-energy density capacitor applications. ACS Appl Mater Interfaces 2014;6:10445–51. doi:10.1021/am502002v.

[46] Treich GM, Tefferi M, Nasreen S, Mannodi-Kanakkithodi A, Li Z, Ramprasad R, et al. A rational co-design approach to the creation of new dielectric polymers with high energy density. IEEE Trans Dielectr Electri Insulat 2017;24:732–43. doi:10.1109/TDEI.2017.006329.

[47] Treich GM, Nasreen S, Mannodi Kanakkithodi A, Ma R, Tefferi M, Flynn J, et al. Optimization of organotin polymers for dielectric applications. ACS Appl Mater Interfaces 2016;8:21270–7. doi:10.1021/acsami.6b04091.

[48] Baldwin AF, Ma R, Mannodi-Kanakkithodi A, Huan TD, Wang C, Tefferi M, et al. Poly(dimethyltin glutarate) as a prospective material for high dielectric applications. Adv Mater 2015;27:346–51. doi:10.1002/adma.201404162.

[49] Baldwin AF, Ma R, Huan TD, Cao Y, Ramprasad R, Sotzing GA. Effect of incorporating aromatic and chiral groups on the dielectric properties of poly(dimethyltin esters). Macromol Rapid Commun 2014;35:2082–8. doi:10.1002/marc.201400507.

[50] Baldwin AF, Huan TD, Ma R, Mannodi-Kanakkithodi A, Tefferi M, Katz N, et al. Rational design of organotin polyesters. Macromolecules 2015;48:2422–8. doi:10.1021/ma502424r.

[51] Nasreen S, Treich GM, Baczkowski ML, Mannodi- Kanakkithodi AK, Baldwin A, Scheirey SK, et al. A material genome approach towards exploration of Zn and Cd coordination complex polyester as dielectrics: design, synthesis and characterization. Polymer (Guildf) 2018;159:95–105. doi:10.1016/j.polymer.2018.10.017.

[52] Li Z, Treich GM, Tefferi M, Wu C, Nasreen S, Scheirey SK, et al. High energy density and high efficiency all-organic polymers with enhanced dipolar polarization. J Mater Chem A Mater 2019;7:15026–30. doi:10.1039/C9TA03601F.

[53] Wu C, Li Z, Treich GM, Tefferi M, Casalini R, Ramprasad R, et al. Dipole-relaxation dynamics in a modified polythiourea with high dielectric constant for energy storage applications. Appl Phys Lett 2019;115. doi:10.1063/1.5123484.

[54] Ma R, Sharma V, Baldwin AF, Tefferi M, Offenbach I, Cakmak M, et al. Rational design and synthesis of polythioureas as capacitor dielectrics. J Mater Chem A Mater 2015;3:14845–52. doi:10.1039/C5TA01252J.

[55] Alamri A, Wu C, Nasreen S, Tran H, Yassin O, Gentile R, et al. High dielectric constant and high breakdown strength polyimide via tin complexation of the polyamide acid precursor. RSC Adv 2022;12:9095–100. doi:10.1039/D1RA06302B.

[56] Lorenzini RG, Greco JA, Birge RR, Sotzing GA. Diels–Alder polysulfones as dielectric materials: computational guidance & synthesis. Polymer (Guildf) 2014;55:3573–8. doi:10.1016/j.polymer.2014.06.041.

[57] Chen L, Kim C, Batra R, Lightstone JP, Wu C, Li Z, et al. Frequency-dependent dielectric constant prediction of polymers using machine learning. NPJ Comput Mater 2020;6:61. doi:10.1038/s41524-020-0333-6.

[58] Wu C, Deshmukh AA, Yassin O, Zhou J, Alamri A, Vellek J, et al. Flexible cyclic-olefin with enhanced dipolar relaxation for harsh condition electrification. Proc Natl Acad Sci 2021;118. doi:10.1073/pnas.2115367118.

[59] Mannodi-Kanakkithodi A, Chandrasekaran A, Kim C, Huan TD, Pilania G, Botu V, et al. Scoping the polymer genome: a roadmap for rational polymer dielectrics design and beyond. Mater Today 2018;21:785–96. doi:10.1016/j.mattod.2017.11.021.

[60] Deshmukh AA, Wu C, Yassin O, Mishra A, Chen L, Alamri A, et al. Flexible polyolefin dielectric by strategic design of organic modules for harsh condition electrification. Energy Environ Sci 2022;15:1307–14. doi:10.1039/D1EE02630E.

[61] Yang L, Yang L, Ma K, Wang Y, Song T, Gong L, et al. Free volume dependence of dielectric behaviour in sandwich-structured high dielectric performances of poly(vinylidene fluoride) composite films. Nanoscale 2021;13:300–10. doi:10.1039/D0NR06070D.

[62] O'Connor KM, Scholsky KM. Free volume effects on the melt viscosity of low molecular weight poly(methyl methacrylate). Polymer (Guildf) 1989;30:461–6. doi:10.1016/0032-3861(89)90014-1.

[63] Low ZX, Budd PM, McKeown NB, Patterson DA. Gas permeation properties, physical aging, and its mitigation in high free volume glassy polymers. Chem Rev 2018;118:5871–911. doi:10.1021/acs.chemrev.7b00629.

[64] Park JY, Paul DR. Correlation and prediction of gas permeability in glassy polymer membrane materials via a modified free volume based group contribution method. J Memb Sci 1997;125:23–39. doi:10.1016/S0376-7388(96)00061-0.

[65] Jansen JC, Macchione M, Tocci E, De Lorenzo L, Yampolskii YP, Sanfirova O, et al. Comparative study of different probing techniques for the analysis of the free volume distribution in amorphous glassy perfluoropolymers. Macromolecules 2009;42:7589–604. doi:10.1021/ma901244d.

[66] An D, Zha X, Gao J, Long Y, Wu K, Li J. Role of free volume on breakdown of epoxy nanocomposites: measurement and manipulation. IEEE Int Conf High Voltage Eng Appl 2022;1–4. doi:10.1109/ICHVE53725.2022.10014459.

[67] Fica-Contreras SM, Li Z, Alamri A, Charnay AP, Pan J, Wu C, et al. Synthetically tunable polymers, free volume element size distributions, and dielectric breakdown field strengths. Mater Today 2023. doi:10.1016/j.mattod.2023.05.010.

[68] Nelson JK, Sabuni H. Evidence for the role of free volume in the high field breakdown of polymers. In: IEEE Conf Electr Insulat Dielectr Phenom; 1980. p. 499–508. doi:10.1109/EIDP.1980.7683883.

[69] Miyamoto T, Shibayama K. Free-volume model for ionic conductivity in polymers. J Appl Phys 1973;44:5372–6. doi:10.1063/1.1662158.

[70] Thomas EM, Nguyen PH, Jones SD, Chabinyc ML, Segalman RA. Electronic, ionic, and mixed conduction in polymeric systems. Annu Rev Mater Res 2021;51:1–20. doi:10.1146/annurev-matsci-080619-110405.

[71] Venkateshvaran D, Nikolka M, Sadhanala A, Lemaur V, Zelazny M, Kepa M, et al. Approaching disorder-free transport in high-mobility conjugated polymers. Nature 2014;515:384–8. doi:10.1038/nature13854.

[72] Singh M, Dong M, Wu W, Nejat R, Tran DK, Pradhan N, et al. Enhanced dielectric strength and capacitive energy density of cyclic polystyrene films. ACS Polym Au 2022;2:324–32. doi:10.1021/acspolymersau.2c00014.

[73] Rabuffi M, Picci G. Status quo and future prospects for metallized polypropylene energy storage capacitors. IEEE Trans on Plasma Sci 2002;30:1939–42. doi:10.1109/TPS.2002.805318.

[74] Zhuang Y, Seong JG, Lee YM. Polyimides containing aliphatic/alicyclic segments in the main chains. Prog Polym Sci 2019;92:35–88. doi:10.1016/j.progpolymsci.2019.01.004.

[75] Gurnani R, Shukla S, Kamal D, Wu C, Hao J, Kuenneth C, et al. AI-assisted discovery of high-temperature dielectrics for energy storage. Nat Commun 2024;15:6107. doi:10.1038/s41467-024-50413-x.

[76] Tran H, Gurnani R, Kim C, Pilania G, Kwon H-K, Lively RP, et al. Design of functional and sustainable polymers assisted by artificial intelligence. Nat Rev Mater 2024. doi:10.1038/s41578-024-00708-8.

[77] Chen L, Pilania G, Batra R, Huan TD, Kim C, Kuenneth C, et al. Polymer informatics: current status and critical next steps. Mater Sci Eng R 2021;144:100595. doi:10.1016/j.mser.2020.100595.

[78] Ellis B, Smith R. Polymers: a property database. CRC Press; 2008. doi:101201/9781420005967.

[79] Wypych G. Handbook of polymers. 2nd ed. ChemTec Publishing; 2016.

[80] Krevelen VDW, Nijenhuis TK. Properties of polymers: their correlation with chemical structure; their numerical estimation and prediction from additive group contributions. 3rd ed. Elsevier Science; 2009.

[81] Braun D, Cherdron H, Rehahn M, Ritter H, Voit B. Polymer synthesis: theory and practice. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. doi:101007/978-3-642-28980-4.

[82] Arndt KF, Lechner MD. Part 2: thermodynamic properties – pVT-Data and thermal properties. Berlin, Heidelberg: Springer Berlin Heidelberg; 2014. doi:101007/978-3-642-41542-5.

[83] Otsuka S, Kuwajima I, Hosoya J, Xu Y, Yamazaki M. PoLyInfo: polymer database for polymeric materials design. In: Int Conf Emerg Intelligent Data Web Tech. IEEE; 2011. p. 22–9. doi:10.1109/EIDWT.2011.13.

[84] Huan TD, Mannodi-Kanakkithodi A, Kim C, Sharma V, Pilania G, Ramprasad R. A polymer dataset for accelerated property prediction and design. Sci Data 2016;3:160012. doi:10.1038/sdata.2016.12.

[85] Huan TD, Ramprasad R. Polymer structure prediction from first principles. J Phys Chem Lett 2020;11:5823–9. doi:10.1021/acs.jpclett.0c01553.

[86] Hayashi Y, Shiomi J, Morikawa J, Yoshida R. RadonPy: automated physical property calculation using all-atom classical molecular dynamics simulations for polymer informatics. NPJ Comput Mater 2022;8:222. doi:10.1038/s41524-022-00906-4.

[87] Phan BK, Shen KH, Gurnani R, Tran H, Lively R, Ramprasad R. Gas permeability, diffusivity, and solubility in polymers: simulation-experiment data

fusion and multi-task machine learning. NPJ Comput Mater 2024;10:186. doi:10.1038/s41524-024-01373-9.

[88] Bicerano J. Prediction of polymer properties. CRC Press; 2002. doi:101201/9780203910115.

[89] Mark JE. Polymer data handbook. 2nd ed. Oxford University Press; 2009.

[90] Sahu H, Shen K-H, Montoya JH, Tran H, Ramprasad R. Polymer structure predictor (PSP): a Python toolkit for predicting atomic-level structural models for a range of Polymer geometries. J Chem Theory Comput 2022;18:2737–48. doi:10.1021/acs.jctc.2c00022.

[91] Rogers D, Hahn M. Extended-connectivity fingerprints. J Chem Inf Model 2010;50:742–54. doi:10.1021/ci100050t.

[92] Tran HD, Kim C, Chen L, Chandrasekaran A, Batra R, Venkatram S, et al. Machine-learning predictions of polymer properties with polymer genome. J Appl Phys 2020;128. doi:10.1063/5.0023759.

[93] Gurnani R, Kuenneth C, Toland A, Ramprasad R. Polymer informatics at scale with multitask graph neural networks. Chem Mater 2023;35:1560–7. doi:10.1021/acs.chemmater.2c02991.

[94] Kuenneth C, Ramprasad R. polyBERT: a chemical language model to enable fully machine-driven ultrafast polymer informatics. Nat Commun 2023;14:4099. doi:10.1038/s41467-023-39868-6.

[95] Kuenneth C, Schertzer W, Ramprasad R. Copolymer informatics with multitask deep neural networks. Macromolecules 2021;54:5957–61. doi:10.1021/acs.macromol.1c00728.

[96] Kuenneth C, Rajan AC, Tran H, Chen L, Kim C, Ramprasad R. Polymer informatics with multi-task learning. Patterns 2021;2:100238. doi:10.1016/j.patter.2021.100238.

[97] Kern J, Chen L, Kim C, Ramprasad R. Design of polymers for energy storage capacitors using machine learning and evolutionary algorithms. J Mater Sci 2021;56:19623–35. doi:10.1007/s10853-021-06520-x.

[98] Kim C, Batra R, Chen L, Tran H, Ramprasad R. Polymer design using genetic algorithm and machine learning. Comput Mater Sci 2021;186:110067. doi:10.1016/j.commatsci.2020.110067.

[99] Batra R, Dai H, Huan TD, Chen L, Kim C, Gutekunst WR, et al. Polymers for extreme conditions designed using syntax-directed variational autoencoders. Chem Mater 2020;32:10489–500. doi:10.1021/acs.chemmater.0c03332.

[100] Gurnani R, Kamal D, Tran H, Sahu H, Scharm K, Ashraf U, et al. polyG2G: a novel machine learning algorithm applied to the generative design of polymer dielectrics. Chem Mater 2021;33:7008–16. doi:10.1021/acs.chemmater.1c02061.

[101] Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J Chem Inf Comput Sci 1988;28:31–6. doi:10.1021/ci00057a005.

[102] RDKit. Open source toolkit for cheminformatics. https://www.rdkit.org/; 2024 [accessed 13 November 2024].

[103] Morgan HL. The generation of a unique machine description for chemical structures-a technique developed at chemical abstracts service. J Chem Doc 1965;5(2):107–13.

[104] Huan TD, Mannodi-Kanakkithodi A, Ramprasad R. Accelerated materials property predictions and design using motif-based fingerprints. Phys Rev B 2015;92:014106. doi:10.1103/PhysRevB.92.014106.

[105] Kuenneth C, Lalonde J, Marrone BL, Iverson CN, Ramprasad R, Pilania G. Bioplastic design using multitask deep neural networks. Commun Mater 2022;3:96. doi:10.1038/s43246-022-00319-2.

[106] Venkatram S, Batra R, Chen L, Kim C, Shelton M, Ramprasad R. Predicting crystallization tendency of polymers using multifidelity information fusion and machine learning. J Phys Chem B 2020;124:6046–54. doi:10.1021/acs.jpcb.0c01865.

[107] Patra A, Batra R, Chandrasekaran A, Kim C, Huan TD, Ramprasad R. A multifidelity information-fusion approach to machine learn and predict polymer bandgap. Comput Mater Sci 2020;172:109286. doi:10.1016/j.commatsci.2019.109286.

[108] Lightstone JP, Chen L, Kim C, Batra R, Ramprasad R. Refractive index prediction models for polymers using machine learning. J Appl Phys 2020;127. doi:10.1063/5.0008026.

[109] Shukla SS, Kuenneth C, Ramprasad R. Polymer informatics beyond homopolymers. MRS Bull 2024;49:17–24. doi:10.1557/s43577-023-00561-0.

[110] Caruana R. Multitask learning. Mach Learn 1997;28:41–75. doi:10.1023/A:1007379606734.

[111] Gilmer J, Schoenholz SS, Riley PF, Vinyals O, Dahl GE. Neural message passing for quantum chemistry. Proc Machine Learn Res 2017;70:1263–72. doi:10.48550/arXiv.1704.01212.

[112] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. CoRR 2017 abs/1706.03762. doi:10.48550/arXiv.1706.03762.

[113] He P., Liu X., Gao J., Chen W. DeBERTa: decoding-enhanced BERT with disentangled attention. AarXiv: 2006, 03654. https://doi.org/10.48550/arXiv.2006.03654.

[114] Wu C, Chen L, Deshmukh A, Kamal D, Li Z, Shetty P, et al. Dielectric polymers tolerant to electric field and temperature extremes: integration of phenomenology, informatics, and experimental validation. ACS Appl Mater Interfaces 2021;13:53416–24. doi:10.1021/acsami.1c11885.

[115] Chen L, Kern J, Lightstone JP, Ramprasad R. Data-assisted polymer retrosynthesis planning. Appl Phys Rev 2021;8. doi:10.1063/5.0052962.

[116] Dai H, Tian Y, Dai B, Skiena S, Song L. Syntax-directed variational autoencoder for structured data. CoRR 2018 abs/1802.08786. doi:10.48550/arXiv.1802.08786.

[117] Huang X, Jiang P, Tanaka T. A review of dielectric polymer composites with high thermal conductivity. IEEE Electr Insul Mag 2011;27:8–16. doi:10.1109/MEI.2011.5954064.

[118] Patel RA, Borca CH, Webb MA. Featurization strategies for polymer sequence or composition design by machine learning. Mol Syst Des Eng 2022;7:661–76. doi:10.1039/D1ME00160D.

[119] Webb MA, Jackson NE, Gil PS, de Pablo JJ. Targeted sequence design within the coarse-grained polymer genome. Sci Adv 2020;6. doi:10.1126/sciadv.abc6216.

[120] Pilania G, Iverson CN, Lookman T, Marrone BL. Machine-learning-based predictive modeling of glass transition temperatures: a case of polyhydroxyalkanoate homopolymers and copolymers. J Chem Inf Model 2019;59:5013–25. doi:10.1021/acs.jcim.9b00807.

[121] Reis M, Gusev F, Taylor NG, Chung SH, Verber MD, Lee YZ, et al. Machine-learning-guided discovery of 19 F MRI agents enabled by automated copolymer synthesis. J Am Chem Soc 2021;143:17677–89. doi:10.1021/jacs.1c08181.

[122] Hanaoka K. Deep Neural Networks for Multicomponent Molecular Systems. ACS Omega 2020;5:21042–53. doi:10.1021/acsomega.0c02599.

[123] Tao L, Byrnes J, Varshney V, Li Y. Machine learning strategies for the structure-property relationship of copolymers. IScience 2022;25:104585. doi:10.1016/j.isci.2022.104585.

[124] Werner M, Guo Y, Baulin VA. Neural network learns physical rules for copolymer translocation through amphiphilic barriers. NPJ Comput Mater 2020;6:72. doi:10.1038/s41524-020-0318-5.

[125] Wheatle BK, Fuentes EF, Lynd NA, Ganesan V. Design of polymer blend electrolytes through a machine learning approach. Macromolecules 2020;53:9449–59. doi:10.1021/acs.macromol.0c01547.

[126] Xu J, Luo T. Unlocking enhanced thermal conductivity in polymer blends through active learning. NPJ Comput Mater 2024;10:74. doi:10.1038/s41524-024-01261-2.

[127] Zhou T, Qiu D, Wu Z, Alberti SAN, Bag S, Schneider J, et al. Compatibilization efficiency of graft copolymers in incompatible polymer blends: dissipative particle dynamics simulations combined with machine learning. Macromolecules 2022;55:7893–907. doi:10.1021/acs.macromol.2c00821.

[128] Sharma A, Mukhopadhyay T, Rangappa SM, Siengchin S, Kushvaha V. Advances in computational intelligence of polymer composite materials: machine learning assisted modeling. Anal Des Arch Comput Methods Eng 2022;29:3341–85. doi:10.1007/s11831-021-09700-9.

[129] Kim S, Schroeder CM, Jackson NE. Open macromolecular genome: generative design of synthetically accessible polymers. ACS Polymers Au 2023;3:318–30. doi:10.1021/acspolymersau.3c00003.

[130] Ohno M, Hayashi Y, Zhang Q, Kaneko Y, Yoshida R. SMiPoly: generation of a synthesizable polymer virtual library using rule-based polymerization reactions. J Chem Inf Model 2023;63:5539–48. doi:10.1021/acs.jcim.3c00329.

[131] Ethier JG, Casukhela RK, Latimer JJ, Jacobsen MD, Rasin B, Gupta MK, et al. Predicting phase behavior of linear polymers in solution using machine learning. Macromolecules 2022;55:2691–702. doi:10.1021/acs.macromol.2c00245.

[132] Kern J, Venkatram S, Banerjee M, Brettmann B, Ramprasad R. Solvent selection for polymers enabled by generalized chemical fingerprinting and machine learning. Phys Chem Chem Phys 2022;24:26547–55. doi:10.1039/D2CP03735A.

[133] Liu T-L, Liu L-Y, Ding F, Li Y-Q. A machine learning study of polymer-solvent interactions. Chin J Polym Sci 2022;40:834–42. doi:10.1007/s10118-022-2716-2.

[134] Nistane J, Chen L, Lee Y, Lively R, Ramprasad R. Estimation of the Flory-Huggins interaction parameter of polymer-solvent mixtures using machine learning. MRS Commun 2022;12:1096–102. doi:10.1557/s43579-022-00237-x.

[135] Bejagam KK, Lalonde J, Iverson CN, Marrone BL, Pilania G. Machine learning for melting temperature predictions and design in polyhydroxyalkanoate-based biopolymers. J Phys Chem B 2022;126:934–45. doi:10.1021/acs.jpcb.1c08354.

[136] Schmid M, Altmann D, Steinbichler G. A simulation-data-based machine learning model for predicting basic parameter settings of the plasticizing process in injection molding. Polymers (Basel) 2021;13:2652. doi:10.3390/polym13162652.

[137] Gansen A, Hennicker J, Sill C, Dheur J, Hale JS, Baller J. Melt instability identification using unsupervised machine learning algorithms. Macromol Mater Eng 2023;308. doi:10.1002/mame.202200628.

[138] Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language models are few-shot learners. CoRR 2020 abs/2005.14165. doi:10.48550/arXiv.2005.14165.

[139] Devlin J, Chang M-W, Lee K, Toutanova KBERT. Pre-training of deep bidirectional transformers for language understanding. CoRR 2018 abs/1810.04805. doi:10.48550/arXiv.1810.04805.

[140] Neethu MS, Rajasree R. Sentiment analysis in twitter using machine learning techniques. In: Fourth Int Conf Comp Comm Network Tech IEEE; 2013. p. 1–5. doi:10.1109/ICCCNT.2013.6726818.

[141] Wu L, Xia Y, Tian F, Zhao L, Qin T, Lai J, et al. Adversarial neural machine translation. CoRR 2018:534–49. doi:10.48550/arXiv.1704.06933.

[142] Tida VS, Hsu SHY. Universal spam detection using transfer learning of BERT model. CoRR 2022 abs/2202.03480. doi:10.48550/arXiv.2202.03480.

[143] Shetty P, Rajan AC, Kuenneth C, Gupta S, Panchumarti LP, Holm L, et al. A general-purpose material property data extraction pipeline from large polymer corpora using natural language processing. NPJ Comput Mater 2023;9:52. doi:10.1038/s41524-023-01003-w.

[144] Bradford G, Lopez J, Ruza J, Stolberg MA, Osterude R, Johnson JA, et al. Chemistry-informed machine learning for polymer electrolyte discovery. ACS Cent Sci 2023;9:206–16. doi:10.1021/acscentsci.2c01123.

[145] Rodriguez F, Cohen F, Ober CK, Archer L. Principles of polymer systems. CRC Press; 2003. doi:101201/b12837.

[146] Zhang C, He X, Lu Q. High-frequency low-dielectric-loss in linear-backbone-structured polyimides with ester groups and ether bonds. Commun Mater 2024;5:55. doi:10.1038/s43246-024-00502-7.

[147] Chen K, Liu Z, Zheng W, Liu S, Chi Z, Xu J, et al. Research progress of intrinsic polymer dielectrics with high permittivity. IET Nanodielectr 2023;6:182–211. doi:10.1049/nde2.12054.

[148] Wei J, Zhu L. Intrinsic polymer dielectrics for high energy density and low loss electric energy storage. Prog Polym Sci 2020;106:101254. doi:10.1016/j.progpolymsci.2020.101254.

[149] Li Z, Wu C, Chen L, Wang Y, Mutulu Z, Uehara H, et al. Probing electronic band structures of dielectric polymers via pre-breakdown conduction. Adv Mater 2024;36:2310497. doi:10.1002/adma.202310497.

[150] Gao L, Lin J, Wang L, Du L. Machine learning-assisted design of advanced polymeric materials. Acc Mater Res 2024;5:571–84. doi:10.1021/accountsmr.3c00288.

[151] De Pablo JJ, Jackson NE, Webb MA, Chen L-Q, Moore JE, Morgan D, et al. New frontiers for the materials genome initiative. NPJ Comput Mater 2019;5:41. doi:10.1038/s41524-019-0173-4.

[152] Huang X, Zhao CY, Wang H, Ju S. AI-assisted inverse design of sequence-ordered high intrinsic thermal conductivity polymers. Mater Today Phys 2024;44:101438. doi:10.1016/j.mtphys.2024.101438.

[153] Dangayach R, Jeong N, Demirel E, Uzal N, Fung V, Chen Y. Machine learning-aided inverse design and discovery of novel polymeric materials for membrane separation. Environ Sci Technol 2025;59:993–1012. doi:10.1021/acs.est.4c08298.

[154] Beaucage PA, Sutherland DR, Martin TB. Automation and machine learning for accelerated polymer characterization and development: past, potential, and a path forward. Macromolecules 2024;57:8661–70. doi:10.1021/acs.macromol.4c01410.

[155] Lu S, Jayaraman A. Machine learning for analyses and automation of structural characterization of polymer materials. Prog Polym Sci 2024;153:101828. doi:10.1016/j.progpolymsci.2024.101828.

[156] Finster R, Sankaran P, Bihar E. Computational and AI-driven design of hydrogels for bioelectronic applications. Adv Electron Mater 2025:2400763. doi:10.1002/aelm.202400763.

[157] Manikandan S, Kaviya RS, Shreeharan DH, Subbaiya R, Vickram S, Karmegam N, et al. Artificial intelligence-driven sustainability: enhancing carbon capture for sustainable development goals–A review. Sus Develop 2024. doi:10.1002/sd.3222.